Zhixiang Yin
Linqiang Pan
Xianwen Fang (Eds.)

Proceedings of The Eighth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA), 2013

Springer

Advances in Intelligent Systems and Computing

Volume 212

Zhixiang Yin · Linqiang Pan
Xianwen Fang
Editors

Proceedings of The Eighth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA), 2013

Editors

Zhixiang Yin
School of Science
Anhui University of Science
   and Technology
Huainan
People's Republic of China

Xianwen Fang
Anhui University of Science
   and Technology
Huainan
People's Republic of China

Linqiang Pan
Huazhong University of Science
   and Technology
Wuhan
People's Republic of China

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)
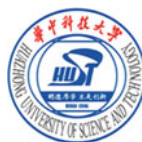
Preface

Bio-Inspired Computing: Theories and Applications (BIC-TA) is one of the flagship conferences on Bio-Computing bringing together the world's leading scientists from different areas of Natural Computing. Since 2006, the conferences have taken place at Wuhan (2006), Zhengzhou (2007), Adelaide (2008), Beijing (2009), Liverpool and Changsha (2010), Malaysia (2011), and India (2012). Following the successes of previous events, The Eighth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA 2013) is being organized and hosted in China by Anhui University of Science and Technology, from 12th to 14th July, 2013. BIC-TA 2013 aims to provide a high-level international forum for researchers with different backgrounds working in the related areas to present their latest results and exchange ideas. The conference has four main sections: Evolutionary computing, Neural computing, DNA computing, and Membrane computing. In order to integrate these sections, the conference organizers invited some experts from different areas of Bio-Inspired Computing to give plenary talks. In this conference, more than 500 conference papers were received while 145 papers were recommended to be published in the Springer book series of Advances in Intelligent Systems and Computing and 60 papers were recommended to be published in SCI Journal, the adopting rate was not more than 30 %. Additionally, the growing trend in Emergent Systems has resulted into the inclusion of two other closely related fields, namely Complex Systems, and Computational Neuroscience, in the BIC-TA 2013 event.

BIC-TA 2013 has attracted a wide spectrum of interesting research papers on various aspects of Bio-Inspired Computing with a diverse range of simulation applications, theories, and techniques within the domain. We much hope that this publication will become an important reference source to many students, researchers, and academics in their educational, research, and professional activities.

The authors are to be commended for their valuable contributions. The editors would like to express their sincere gratitude to the reviewers, track chairs, and program committee members, who have done justice to the entire review process and have helped to maintain the quality and clarity of presentation of the papers.

We would like to acknowledge the members of the BIC-TA 2013 organizing committee for their efforts in organizing the conference. The organizing committee

benefited from the support it received from the School of Anhui University of Science and Technology. We would also like to thank the members of the BIC-TA steering committee, especially Linqiang Pan and Xu Jin, for their guidance and advice. We are indebted to the members of the BIC-TA program committee and additional reviewers for their diligent and careful reviewing which led to valuable improvements in the accepted papers. Finally, we would like to thank all the presenters and authors for their active participation at BIC-TA 2013, which made the conference a success. Special thanks are given to Springer-Verlag for his encouragement and help to our work. We would like to acknowledge here once again all the college and co-workers who contributed to the success of this interesting and stimulating conference.

It is envisaged that the BIC-TA conference series will continue to grow and include relevant future research and development challenges in this field of Computing.

Huainan, Anhui, January 2013   Xianwen Fang

Contents

Part I
Theories

Remarks on Invariant Method of the Second-Order Linear Differential Equations with Variable Coefficients

Linlong Zhao

Abstract Invariant method was used to solve the linear second-order equations with variable coefficients. We employ the invariant variable method to give the integrable condition of equations and to display the superiority of this method.

1 Introduction

In the second order liner differential equations,

$$y'' + p(x)y' + q(x)y = f(x) \tag{1}$$

As we have already known function p(x) and q(x) with special condition, the Eq. (1) just can be solved with the elementary solution method. Therefore, it is decisive job to look for the new relation for p(x) and q(x) to make the original Eq. (1) become solvable equation. Recently, many authors have worked in this field [1–6]. This method that looking for p(x) and q(x) has no general regulation. So it becomes difficulty to give new relation.

Now, we make use of the invariant argument of the second order linear differential equations to make a discussion on these problems again.

L. Zhao (✉)
Department of Mathematics, Ankang University, 92 Yucai Road, Ankang 725000,
People's Republic of China
e-mail: aktczll@163.com

2 The Invariant Variable of the Second Order Linear Differential Equations

For Eq. (1), let the following transformation

$$y = z(x)e^{\int \varphi(x)dx} \tag{2}$$

where $\varphi(x)$ is treat to settle function. We can obtain equation as follows

$$z'' + A(x)z' + B(x)z = f(x)e^{-\int \varphi(x)dx} \tag{3}$$

where

$$A(x) = 2\varphi(x) + p(x) \tag{4}$$

$$B(x) = \varphi'(x) + \varphi^2(x) + p(x)\varphi(x) + q(x) \tag{5}$$

Employing the conditions (4) and (5), we get equivalence relation,

$$2p'(x) + p^2(x) - 4q(x) = 2A'(x) + A^2(x) - 4B(x) \tag{6}$$

$$\text{So}, I = 2p'(x) + p^2(x) - 4q(x) \tag{7}$$

which is defined as invariant argument of Eq. (1).

3 The Invariant Method of the Second Order Liner Differential Equations

In 1998, Zhao Linlong [7, 8] obtained the invariant method of the second order linear differential equations.

Theorem 1 [7] *For Eq. (1), if there is differential function $\varphi(x)$ satisfying the following condition,*

$$I = 2p'(x) + p^2(x) - 4q(x) = 2\varphi'(x) + \varphi^2(x) \tag{8}$$

then Eq. (1) has the solution as follows,

$$y = z(x)e^{\frac{1}{2}\int (\varphi(x) - p(x))dx} \tag{9}$$

where $z(x)$ satisfying the following linear differential equations,

$$W' + \varphi(x)W = f(x)e^{-\frac{1}{2}\int (\varphi(x) - p(x))dx} \ (W = z') \tag{10}$$

4 The Superiority of Invariant Method for the Second Order Liner Differential Equations

Obviously, whether the invariant variable method of second order linear differential equations can be used is completely determined by the existence of the function $\varphi(x)$ of Eq. (8). This method make look for relation $p(x), q(x)$ with regulation to solve the function $\varphi(x)$ in the Eq. (8) instead of looking for this relation with no regulation. This method is simple and has the following advantages.

4.1 Connection Solvable Method of Second Order Linear Differential Equation Constant Coefficient with Variable Coefficient

In Eq. (1), if $p(x), q(x)$ are constants, then differential function $\varphi(x)$ for Eq. (8) satisfies the following form

$$\varphi(x) = \pm \sqrt{I} \left(I = p^2 - 4q \right) \tag{11}$$

Corollary 1 *If the invariant variable is a constant coefficient I for Eq. (8), then the solution to Eq. (1) has the following form*

$$y = z(x)e^{\frac{1}{2}\int (\sqrt{I} - p(x))dx} \tag{12}$$

where $I = 2p'(x) + p^2(x) - 4q(x)$, and $z(x)$ satisfies one order linear differential equation

$$W' + \sqrt{I}W = f(x)e^{-\frac{1}{2}\int (\sqrt{I} - p(x))dx} \left(W = z', I = 2p'(x) + p^2(x) - 4q(x) \right) \tag{13}$$

4.2 Connection Second Order Linear Homogeneous Differential Equation and Non-Homogeneous Equation

In Eq. (1), if $f(x)=0$, then Eq. (10) can be changed into the following form

$$W' + \varphi(x)W = 0 \tag{14}$$

where $W = z'$, $\varphi(x)$ satisfying Eq. (8) .

Corollary 2 *In the second order Euler ordinary differential equations*

$$x^2 y'' + axy' + by = f(x)\,(a,b) \text{ is coefficients} \tag{15}$$

If there is a constant coefficient k satisfying

$$k^2 - 2k - a^2 + 2a + 4b = 0 \tag{16}$$

then Eq. (1) has the solution

$$y = z(x)e^{\frac{1}{2}\int (\frac{k}{x} - p(x))dx} = x^{\frac{k}{2}} z(x) e^{-\frac{1}{2}\int p(x)dx} \tag{17}$$

where $z(x)$ satisfies the one order linear ordinary differential equation

$$W' + \frac{k}{x} W = x^{-\frac{k}{2}-2} f(x) e^{\frac{1}{2}\int p(x)dx}\ (W = z') \tag{18}$$

According to the relation of invariant variable $I = 2p'(x) + p^2(x) - 4q(x) = \frac{-2a}{x^2} + \frac{a^2}{x^2} - \frac{4b}{x^2}$,

Let $\varphi(x) = \frac{k}{x}$ (k is coefficient, then $I = 2\varphi'(x) + \varphi^2(x) = \frac{-2k}{x^2} + \frac{k^2}{x^2}$, we can get Eqs. (16) and (18).

4.3 Connection Second Order Linear Differential Equation the General Solution with Special Solution

According to Eq. (10), we will get

$$z' = \left(\int f(x)e^{\frac{1}{2}\int (p(x)-\varphi(x))dx} dx + c\right)e^{-\int (\varphi(x)dx} \tag{19}$$

$$z = \int \left(\int f(x)e^{\frac{1}{2}\int (p(x)-\varphi(x))dx} dx + c_1\right)e^{-\int (\varphi(x)dx} dx + c_2 \tag{20}$$

then Eq. (1) has the general solution

$$y = \left(\int \left(\int f(x)e^{\frac{1}{2}\int (p(x)-\varphi(x))dx} dx + c_1\right)e^{-\int (\varphi(x)dx} dx + c_2\right)e^{\frac{1}{2}\int (\varphi(x)-p(x))dx} \tag{21}$$

When $c_1 = c_2 = 0$, Eq. (1) has the special solution

$$y = \left(\int \left(\int f(x)e^{\frac{1}{2}\int (p(x)-\varphi(x))dx} dx\right)e^{-\int (\varphi(x)dx} dx\right)e^{\frac{1}{2}\int (\varphi(x)-p(x))dx} \tag{22}$$

Corollary 3 *If there is differential function $\varphi(x)$ for Eq. (1) satisfying (8), then there is the special solution of Eq. (22)*

$$\int (\int f(x)e^{\frac{1}{2}\int (p(x)-\varphi(x))dx}dx)e^{-\int (\varphi(x)dx}dx \qquad (23)$$

In coefficient linear differential Eq. (1), this expands the special solution of function $f(x)$ form,

$$f(x) = (A_m(x)\cos \beta x + B_n(x)\sin \beta x)e^{\alpha x} \qquad (24)$$

where $A_m(x), B_n(x)$ is polynomial, α, β is real.

5 Discussion About invariant Method of the Second Order Linear Differential Equations

5.1 The Integrable of the Second-Order Linear Differential Equation

In 1841, Liouville proved Riccati equation

$$y' = y^2 + x^2 \qquad (25)$$

which has not elementary solution. But by transformation $y = \frac{z'}{-z}$, then it become another no elementary solution equation

$$z'' + x^2 z = 0 \qquad (26)$$

5.2 Discussion for Integrable Condition of the Second Order Linear Differential Equations

In   Eq.(8),   Let $\varphi(x) = p(x) + 2\mu(x)\sqrt{\varepsilon q(x)}$, $\mu'(x) = -\sqrt{\varepsilon q(x)}(\mu^2(x) + \frac{a}{2}\mu(x) + 1)$,
when $q(x) > 0$, $\varepsilon = 1$ or $q(x) < 0$, $\varepsilon = -1$, a is constant, we can obtain

$$\frac{q'(x) + 2p(x)q(x)}{q^{\frac{3}{2}}(x)} = 2a \qquad (27)$$

Theorem 2[3] *If Eq. (1) satisfies (27), then it can be transformed into the following equation with constant coefficient*

$$\varepsilon \frac{d^2 y}{dt^2} + \frac{a}{2} \frac{dy}{dt} + y = \frac{f(x)}{q(x)} \quad (q(x) > 0, \varepsilon = 1 \ or \ q(x) < 0, \varepsilon = -1) \qquad (28)$$

In 2000, it was given in [8]. It is clear that the method invariant variable is general and need not to look integrable type of Eq. (1) avoiding the previous integrable type.

6 Applications

Example 1 [3] *Solve an equation* $y'' + \frac{2}{x} y' + y = \frac{1}{x} \cos x$, *where* $r(x) = \frac{\sin x}{x}$ *is special solution of equation.*

Solve: $I = 2(\frac{2}{x})' + (\frac{2}{x})^2 - 4 = -4$. *By Corollary 1, we has the equation,*

$$W' + 2iW = \frac{1}{x} \cos x e^{-\frac{1}{2} \int (2i - \frac{2}{x}) dx} = \cos x e^{-ix}$$

$$z' = W = (\int \cos x e^{ix} dx + c) e^{-2ix} = (\int (\cos^2 x + i \cos x \sin x) dx + c) e^{-2ix}$$

$$= \frac{x}{2} e^{-2ix} - \frac{i}{4} + c e^{-2ix}$$

$$z = \int (\frac{x}{2} e^{-2ix} - \frac{i}{4} + c e^{-2ix}) dx = \frac{ix}{4} e^{-2ix} + \frac{1}{8} e^{-2ix} - \frac{i}{4} x + c_1 e^{-2ix} + c_2$$

$$= \frac{ix}{4} e^{-2ix} - \frac{i}{4} x + c_1 e^{-2ix} + c_2$$

Then we obtain the general solution of equation,

$$y = z(x) e^{\frac{1}{2} \int (2i - \frac{2}{x}) dx} = \frac{i}{4} e^{-ix} - \frac{i}{4} e^{ix} + \frac{c_1}{x} e^{-ix} + \frac{c_2}{x} e^{ix} = \frac{\sin x}{2} + \frac{c_1}{x} e^{-ix} + \frac{c_2}{x} e^{ix}$$

The new method for solving the linear second-order equations displays the superiority that it solves method without the special solution of equation.

Example 2 [2] *Solve an equation* $y'' - \frac{2}{x} y' + \frac{2}{x^2} y = x(\cos x - x \sin x)$.

Solve: *In Euler differential equation, by Corollary 2, has the equation* $k^2 - 2k - 4 - 4 + 8 = k^2 - 2k = 0$, *Let* $k = 0$, *then has the equation*

$$W' = x^{-2} f(x) e^{\frac{1}{2} \int p(x) dx} = x(\cos x - x \sin x) e^{-\int \frac{1}{x} dx} = \cos x - x \sin x \ (W = z')$$

$$w = z' = x \cos x + c, z = x \sin x - \cos x + c_1 x + c_2$$

$$y = z(x)e^{-\frac{1}{2}\int p(x)dx} = (x\sin x - \cos x + c_1 x + c_2)e^{\int \frac{1}{x}dx}$$
$$= x(x\sin x - \cos x + c_1 x + c_2)$$

Example 3 [6] *Solve an equation* $y'' - 4y' + 4y = (3x^2 - 2x + 1)e^{2x}$.

Solve: *Since* $I = 16 - 16 = 0$. *By corollary 1, we has the equation*

$$W' = f(x)e^{\frac{1}{2}\int p(x)dx} = (3x^2 - 2x + 1)e^{2x}e^{-2x} = 3x^2 - 2x + 1$$

$$Z' = W = x^3 - x^2 + x + c, \ z = \frac{1}{4}x^4 - \frac{1}{3}x^3 + \frac{1}{2}x^2 + c_1 x + c_2$$

Then we has the general solution of equation

$$y = z(x)e^{-\frac{1}{2}\int p(x)dx} = (\frac{1}{4}x^4 - \frac{1}{3}x^3 + \frac{1}{2}x^2 + c_1 x + c_2)e^{2x}$$

References

1. Deng Y (2011) The general solution to construct on two classes of second order linear homogeneous differential equation for variable coefficient. J Southwest Chin Normal Univ (Nat Sci Ed). 6:1–5
2. Luxiang F (2011) General solution of second order nonlinear differential equations with variable coefficients and its application. J Chin Three Gorges Univ (Nat Sci) 5:96–98
3. Xiaoqin FAN, Zhaohui BI (2011) Several kinds of second-order variable coefficient differential equation solution. Coll Math 3:200–203
4. Xiao LP, Chen ZX (2011) Some properties of solutions of periodic second order linear differential equations. J Math Res Exposition 2:279–286
5. Xiong C, Xie J (2010) Simplification of solution of differential equation with second-order constant coefficients. J Nanchang Inst Technol 1:5–8
6. Chen C, Shen J, Ding S (2009) A new solution of linear differential equations with constant coefficient. J Chongqing Jiaotong Univ (Nat Sci) 1:158–160
7. Linlong Z (1998) A method for general solution the order 2 linear differential equations. J Changsha Univ 2:5–8
8. Linlon Z (2000) Often the differential equation studies a new theory. Map publisher in Xi'an, Xi'an

Convergence Analysis on Immune Optimization Solving CCP Problems

Zhuhong Zhang and Fei Long

Abstract This work concentrates on studying the property of convergence of a sample allocation-based immune optimization approach used in solving linear or nonlinear chance-constrained programming (CCP) with general random variables. First, we make some theoretical studies about existence of optimal reliable solutions and give an approximate relation between the true CCP and the sample average approximation problem, depending on some statistic and analysis theory. Second, a bio-inspired immune optimization approach is developed to assume solving CCP problems. Our theoretical analysis shows that such approach, which is capable of being formulated by a non-homogeneous Markov model, is convergent. Experimentally, performance searching curves reveal that the approach can obtain valuable performances including the optimized quality, noisy suppression and convergence.

Keywords Chance-constrained programming · Immune optimization · Sample average approximation · Convergence

1 Introduction

In practical optimization problems, objective functions and constraints are often inevitably perturbed by uncertainty, and hence a search procedure becomes extremely difficult when searching for their solutions [1]. CCP is a kind of

Z. Zhang (✉)
Institute of System Science and Information Technology, College of Science, Guizhou University, Guiyang 550025 Guizhou, China
e-mail: sci.zhzhang@gzu.edu.cn

F. Long
College of Computer Science and Information, Guizhou University, Guiyang 550025 Guizhou, China
e-mail: flong1973@yahoo.com.cn

stochastic programming with at least a chance or probabilistic one. The key of solving it is to study efficient approximation or transformation approaches to handle chance constraints. In comparison with mathematical approaches for CCP, intelligent optimization is a more useful tool because of simplicity and effectiveness, in which stochastic simulation [2] is often used to deal with uncertainties. Even if a wide range of engineering applications [3], CCP is still studied scarcely by intelligent scholars. The main difficulty is that the noisy environment influences seriously the optimized quality and individual's evaluation, and hence it is almost impossible to obtain the theoretical optimum. Fortunately, several methodological achievements (e.g., see [2, 4]) are reported to solve several kinds of general CCP problems, but their theoretical study, e.g., convergence, is still open. We note that Nakama [5] made a series of meaningful contributions to probing into convergence properties of genetic algorithms (GAs) used for solving expected value optimization problems. Chen and He [6] developed a heuristic sequential allocation procedure to identify the best of individuals in noisy environments. Such procedure as an optimal selection scheme can theoretically find a locally optimal solution asymptotically, but is difficult if applied to CCP problems. Up to now, less theoretical work has been reported about CCP using intelligent optimization.

Immune optimization has become increasingly popular and appeared lots of valuable achievements [7]. To our knowledge, in such branch we take a first step to investigate a bio-inspired immune optimization approach for general CCP problems [8], whereas its theoretical foundations keep open. In the present work, we first study an approximate relation between the CCP problem and the approximation problem; we next develop a general optimization mechanism, i.e., sample allocation-based immune optimization approach (SAIOA), which bases on the above approach. Finally, SAIOA's convergence property is studied by means of Markov theory and the law of large numbers.

2 Problem Description and Existence of Solutions

Consider the following general chance-constrained programming problem (P_α):

$$\underset{x \in D}{Min} f(x) = E[F(x, \xi)]$$
$$s.t. \begin{cases} p(x) = \Pr\{G(x, \xi) \leq 0\} \geq 1 - \alpha, \\ \quad g(x) \leq 0, h(x) = 0, \end{cases}$$

with bounded and closed domain D in R^p, decision vector x in D, random vector ξ in Ω with $\Omega \subseteq R^q$ and significance level α in the interval [0,1], where E[.] and Pr{.} are the operators of expectation and probability respectively; $F(x, \xi)$ and $G(x, \xi)$ denote the general linear or nonlinear stochastic objective and constraint functions, respectively; $g(x)$ and $h(x)$ are the deterministic vector-valued constraint functions with J and K dimensions respectively. We prescribe that the symbol of

$x \leq 0$ stands for $x_i \leq 0$, $i = 1, 2, ..., p$, with $x = (x_1, x_2 ,..., x_p)$. Additionally, once a joint chance constraint appears in practical problems, e.g., $\Pr\{G_i(x, \xi) \leq 0, i = 1, 2, ..., K\} \geq 1 - \alpha$, it can be equivalently transformed into the above chance constraint by $G(x, \xi) = \max\{G_i(x, \xi), i = 1, 2, ..., K\}$. We say that x is reliable with significance level α or that x is a reliable candidate, if it satisfies the above constraints; otherwise, it is called unreliable. All such reliable candidates constitute the feasible region D_α, where we always assume $D_\alpha \neq \varphi$ in the following theoretical studies. A reliable candidate is called an optimal reliable solution if it possesses the minimal objective value among reliable candidate solutions. All such optimal reliable solutions form a set D_α^*, and the minimum is expressed by z_α^*. Let $\Gamma(x)$, which measures the degree that x is far from the feasible region, be the quantity of constraint violation for x (e.g., see [8]). We next investigate some properties of P_α, based on the following basic assumptions.

(H1) $F(x, \xi)$ is continuous in x and measurable in ξ; there exists an integrable function $\gamma: \Omega \rightarrow R^+$ satisfying $| F(x, \xi)| \leq \gamma(\xi)$ for all $x \in D$;

(H2) $G(x, \xi)$, $g(x)$ and $h(x)$ are continuous in x, and $G(x, \xi)$ is measurable in ξ;

(H3) $p(x^*) > 1 - \alpha$ for each $x^* \in D_\alpha^*$.

The above assumptions guarantee that P_α has the following property:

Theorem 1 *If (H1) and (H2) hold, P_α has at least an optimal reliable solution.*

Proof According to (H1), we claim that $f(x)$ is continuous in x. To this end, suppose that x_n converges to $\bar{x}$ when $n \rightarrow \infty$ with any $x_n, \bar{x} \in D$. It follows from (H1) and Lebesgue dominated convergence theorem that

$$\lim_{n \to \infty} f(x_n) = \int_\Omega \lim_{n \to \infty} F(x_n, \xi) \, dP(\xi) = f(\bar{x}). \tag{1}$$

This implies that $f(x)$ is continuous. Subsequently, introduce an indicator function

$$I_{(-\infty,0)}(z) = \begin{cases} 1, & if \ z \leq 0, \\ 0, & otherwise. \end{cases} \tag{2}$$

Write $\Psi(x, \xi) = I_{(-\infty,0)}(G(x, \xi))$. Since $I_{(-\infty,0)}(.)$ is upper semicontinuous, by means of (H1) we acquire

$$\Psi(\bar{x}, \xi) \geq \overline{\lim_{G(x,\xi) \to G(\bar{x},\xi)}} I_{(-\infty,0)}(G(x, \xi)) = \overline{\lim_{x \to \bar{x}}} \Psi(x, \xi); \tag{3}$$

namely, $\Psi(x, \xi)$ is upper semicontinuous in x. Hence, Fatou's lemma and Eq. (3) hint that $E[\Psi(x, \xi)]$ is upper semicontinuous. Again,

$$E[\Psi(x, \xi)] = \int_\Omega \Psi(x, \xi) dP(\xi) = \int_{G(x,\xi) \leq 0} dP(\xi) = p(x), \tag{4}$$

this way, p(x) is upper semicontinuous. Hence, as related to (H$_2$), we obtain that D_α is a nonempty compact set. Thus, $f(x)$ can achieve the minimum on D.

3 Sample Average Approximation and Theoretical Analysis

Let $m(x)$ be the sample size of ξ at the point x, and ξ_1, ξ_2, ..., $\xi_{m(x)}$ be $m(x)$ replications of ξ. $p_{m(x)}(x)$ stands for the approximate estimates of $p(x)$ with

$$p_{m(x)}(x) = m(x)^{-1} \sum_{i=1}^{m(x)} I_{(-\infty,0)}(G(x, \xi_i)). \tag{5}$$

So, the sample average approximation problem ($P_\alpha{}^m$) can be formulated:

$$\underset{x \in D}{Min}\ f_{m(x)}(x) = \mu_{m(x)}(F)$$

$$s.t., p_{m(x)}(x) \geq 1 - \alpha, g(x) \leq 0, h(x) = 0,$$

where $\mu_{m(x)}(F)$ denotes the average of $m(x)$ objective observations at the point x. Meanwhile, after all candidates in D are attached sample sizes, $x \in D$ is said to be empirically feasible, if x satisfies the above constraints. Notice that if all these candidates are attached the same sample size, P_α^m is a conventional sample approximation problem of P_α. In the literature, Pagnoncelli et al. [9] investigated the approximate relation between the expected value optimization problem with a bounded variable constraint and the sample average approximation problem. We extend their result to the case where the relation between P_α and P_α^m is developed. Since we expect that the minimum of P_α^m is sufficiently close to that of P_α in the case where $m(x)$ is sufficiently large for each $x \in D$, in the following theoretical analysis we assume that all $m(x)$, $x \in D$, equal the same large integer M. So, P_α^m is expressed by P_α^M. Let D_α^M represent the set of all feasible candidates for P_α^M and $\hat{D}_\alpha^M$ the set of optimal solutions for such problem. Write $z_\alpha^M = \min\{f_M(x), x \in D_\alpha^M\}$. We cite the classical concept of epi-convergence, namely, let $g_n(.)$ and $g(.)$ be real-valued functions from D to R with $n \geq 1$; we say that $\{g_n(.)\}$ epi-converges to $g(.)$, if (1) $g(\bar{x}) \leq \underset{n \to \infty}{\underline{\lim}}\, g_n(x_n)$ whenever $x_n \to \bar{x}$ with $x_n, \bar{x} \in D$, and (2) $\underset{n \to \infty}{\lim} g_n(y_n) = g(\bar{x})$ for at least one sequence $y_n \to \bar{x}$.

Theorem 2 *If the above three assumptions hold, then $z_\alpha^M \to z_\alpha^*$, w.p.l, as $M \to \infty$.*

Proof By means of the above theorem, one can see that D_α^* is nonempty. Further, as associated to (H$_3$), we take $x^* \in D_\alpha^*$, and hence the (strong) law of large numbers yields that $p_M(x^*) \to p(x^*)$, w.p.l, as M $\to \infty$. Thus, we have that $p_M(x^*) \geq 1 - \alpha$ when M is large enough, because $p(x^*) > 1 - \alpha$. This hints that $x^* \in D_\alpha^M$, and accordingly $z_\alpha^M \leq f_M(x^*)$. So, it follows again from the (strong) law of large numbers that

$$\overline{\lim_{n\to\infty}} z_\alpha^M \leq \overline{\lim_{n\to\infty}} f_M(x^*) = z_\alpha^*, w.p.l. \tag{6}$$

In addition, as related to the property of continuity of $\Psi(x, \xi)$ in x, $p_M(x)$ is upper semicontinuous in x. This, along with the assumptions of $g(x)$ and $h(x)$, implies that D_α^M is a compact set. Thus, it follows from continuity of $f_M(x)$ in x and the compactness of D_α^M that P_α has at least an optimal solution when M is sufficiently large. Take $x_M \in \hat{D}_\alpha^M$. Since D is compact, we may suppose that $x_M \to \bar{x}$, $\bar{x} \in D$, as $M \to \infty$, in order to be simplicity of notation. Again, since $1 - p_M(.)$ is lower semicontinuous, we gain that it epi-converges to $1 - p(.)$, $w.p.l.$ This hints

$$1 - p(\bar{x}) \leq \underline{\lim_{n\to\infty}} (1 - p_M(x_M)), \tag{7}$$

and hence

$$p(\bar{x}) \geq \overline{\lim_{M\to\infty}} p_M(x_M) \geq 1 - \alpha. \tag{8}$$

Thus, Eq. (8) and (H$_2$) result in $\bar{x} \in D_\alpha$. This illustrates that $f(\bar{x}) \geq z_\alpha^*$. Again, since $f_M(.)$ is continuous, it follows from (H$_1$) that $f_M(.)$ epi-converges to $f(.)$, $w.p.l$, as $M \to \infty$. Accordingly, we obtain

$$z_\alpha^* \leq f(\bar{x}) \leq \underline{\lim_{n\to\infty}} f_M(x_M) = \underline{\lim_{n\to\infty}} z_\alpha^M, w.p.l. \tag{9}$$

Summarily, Eqs. (6) and (9) ensure that the conclusion is true.

4 Algorithm Statement and Illustrations

We develop a general immune optimization framework for CCP problems, which bases on our previous work [8]. As associated to P_α and $P_\alpha{}^m$, a real-encoded antibody is viewed as a candidate in $P_\alpha{}^m$; the antigen is regarded as P_α. Our task is to find the best antibody (i.e., optimal reliable solution) through running SAIOA on P_α^m. Note that we recall that all candidates in P_α^m have their respective sample sizes under a given sample allocation rule. Let M_n, which increases with iteration number n, denote the sample size of a given antibody population X with size N, namely the total of sample sizes attached by elements in X is equal to M_n. Further, for $x \in X$, $aff(x)$, $\delta(x)$ and $\Gamma_{m(x)}(x)$ represent in order the antibody affinity, suppression radius and constraint violation (i.e., the estimates of $\Gamma(x)$). They can be used to seek diverse antibodies in X based on a niche-like method. To this point, we firstly rank decreasingly all antibodies in X according to their affinities; secondly, those elements, whose affinities are between $aff(x)$ and $aff(x) + \delta(x)$, are said to be suppressed by x. This way, survival antibody x can obtain the number of

antibodies suppressed by it. This number is called the suppression size of x. Based on these preliminaries, we describe SAIOA below.

Step 1. Set $n \leftarrow 1$. Generate initial population A_n of N random antibodies, and set $m(x) = m_0$ for all $x \in A_n$; calculate $f_{m(x)}(x)$ and $\Gamma_{m(x)}(x)$ with all $x \in A_n$.
Step 2. Allocate population sample size M_n to all elements in A_n according to the rule that better antibodies can obtain larger sample sizes.
Step 3. Calculate $f_{m(x)}(x)$, $\delta(x)$ and $\Gamma_{m(x)}(x)$ with all $x \in A_n$.
Step 4. A reproduction scheme is enforced on A_n; namely, those survival antibodies are only admitted to proliferate their clones in terms of their suppression sizes, and hence a clonal population B_n is formed with size N.
Step 5. All clones, with time-varying mutation rates conversely proportional to their respective affinities, are mutated through a mutation rule with ergodicity; such mutated clones with the same sample size m_n constitute a temporary population C_n; subsequently, their empirical averages and constraint violations are calculated.
Step 6. Combine A_n with C_n, and select N better antibodies to form A_{n+1}.
Step 7. If a stopping criterion is satisfied, the procedure is terminated; otherwise, set $n \leftarrow n + 1$, and return Step 2.

5 Convergence Study

SAIOA can be considered as an evolution chain: $A_n \rightarrow B_n \rightarrow C_n \rightarrow A_{n+1}$. Through the algorithm description, A_{n+1} only depends on the state of A_n, while the mutation rate depends on n. So, $\{A_n\}_{n \geq 1}$ is a nonhomogeneous Markov chain. Assume that the decision domain D as in Sect. 3 is a finite set. Let S represent the antibody space; S^N stands for a state space composed of antibody populations with sizes N; $X \in S^N$ is called a state. $\Pr\{x \rightarrow y\}$ presents the probability which x is mutated into y with $x, y \in S$. As associated to SAIOA's formulation, we acquire the following properties.

Lemma 5.1 *If the above assumptions (H_1)–(H_3) holds, and both M_n and m_n are sufficiently large as $n \rightarrow \infty$, then there exists $N_1 > 0$ such that when $n > N_1$,* $\Pr\{A_{n+1} \cap D_\alpha^* = \phi | A_n \cap D_\alpha^* = \phi\} = 0$.

Proof We know that $D_\alpha^* \neq \varphi$ by Theorem 1. Assume that $A_n \cap D_\alpha^* \neq \varphi$. Through Step 3 as in SAIOA, the best antibody in A_n can obtain the largest sample size decided by M_n. Again, since M_n increases with n, it follows from the law of large numbers and (H_3) that there exists a sufficiently large positive integer N_1, such that when $n > N_1$, $p_{m(x^*)}(x^*) \geq 1 - \alpha$ for each $x^* \in A_n \cap D_\alpha^*$. Accordingly, x^* is empirically feasible. We claim that when n is sufficiently large, $aff(x^*) \geq aff(x)$ for any $x \in A_n$. Otherwise, if there exists $x_0 \in A_n$ but $x_0 \notin D_\alpha^*$ such that

$aff(x^*) < aff(x_0)$. We see that x_0 is empirically feasible according to the design rule of antibody affinity. Therefore, $f_{m(x0)}(x_0) < f_{m(x*)}(x^*)$. Again, as related to the sample allocation rule in step 3, $m(x_0)$ must be sufficiently large when $n > N_1$. This, together with the law of large numbers, implies that $f(x_0) \leq f(x^*)$. This is not true because of $x_0 \notin D_\alpha^*$. All these illustrate that x^* is best in A_n when $n > N_1$. On the other hand, if there exists some element x in C_n such that x is not inferior to x^*, x must be empirically feasible for $P_\alpha m$ when $n > N_1$, and hence x must be an optimal reliable solution of P_α. Summarily, when $n > N_1$, we have that $A_{n+1} \cap D_\alpha^* \neq \varphi$ if $A_n \cap D_\alpha^* \neq \varphi$.

Lemma 1 *If there exists $0 < \delta < 1$, such that $Pr\{x \rightarrow y\} \geq \delta$ for any $x, y \in S$, then there exists $0 < \varepsilon < 1$, $0 < \beta < 1$ and $N_2 > 0$, such that $Pr\{A_{n+1} \cap D_\alpha^* \neq \varphi | A_n \cap D_\alpha^* = \varphi\} \geq (1 - \varepsilon^n)^2 \beta$ with $n > N_2$.*

Proof Let $A_n = X$ and $A_{n+1} = Y$, $X, Y \in S^N$, with $X \cap D_\alpha^* = \varphi$ and $Y \subseteq D_\alpha^*$. Based on the formulation of SAIOA, step 3 causes that the empirical values of antibodies in A_n are updated; step 4 selects deterministically some parents to proliferate N clones; step 6 picks up deterministically N better antibodies to constitute the next population from $A_n \cup C_n$ in terms of their empirical values. Thus, the law of large numbers implies that when n is sufficiently large, we can acquire that, (1) $Pr\{B_n = Z_0 | A_n = X\} > 1 - \varepsilon^n$ for some state $Z_0 \subseteq X$, $Z_0 = (z_1, z_2, ..., z_N)$, and (2) $Pr\{A_{n+1} = Y | C_n = Y\} > 1 - \varepsilon^n$, where ε is a sufficiently small positive number. Again, relying upon the assumption of ergodicity, for any $W = (w_1, w_2, ..., w_N) \in S^N$ we obtain $Pr\{z_i \rightarrow w_i\} \geq \delta$, with $1 \leq i \leq N$. Consequently, we derive out that $Pr\{C_n = W | B_n = Z_0\} \geq \delta^N \equiv \beta$. Hence, it derives from K–C equation that

$$Pr\{A_{n+1} = Y | A_n = X\} = \sum_{Z,W \in S^N} Pr\{B_n = Z | A_n = X\}.$$
$$Pr\{C_n = W | B_n = Z\} Pr\{A_{n+1} = Y | C_n = W\} \tag{10}$$
$$> (1 - \varepsilon^n)^2 Pr\{C_n = Y | B_n = Z_0\} \geq (1 - \varepsilon^n)^2 \beta$$

Thereby, there exists a positive integer $N_2 > 0$ such that when $n > N_2$,

$$Pr\{A_{n+1} \cap D_\alpha^* \neq \varphi | A_n \cap D_\alpha^* = \varphi\} \geq Pr\{A_{n+1} = Y | A_n = X\} \geq (1 - \varepsilon^n)^2 \beta. \tag{11}$$

Theorem 3 *If the above three assumptions hold, SAIOA is weakly convergent in probability, i.e., $\lim_{n \to \infty} Pr\{A_n \cap D_\alpha^* \neq \varphi\} = 1$.*

Proof Since $0 < \varepsilon < 1$ and $0 < \beta < 1$, there exists $N_3 > 0$ and $0 < \eta < 1$ such that when $n > N_3$, we have $\eta \geq 1 - (1 - \varepsilon^{N_3})$. Take $N_0 = \max\{N_1, N_2, N_3\}$. When $n > N_0$, Lemmas 1 and 2 indicate that

$$Pr\{A_{n+1} \cap D_\alpha^* = \varphi\} \leq (1 - \eta) Pr\{A_n \cap D_\alpha^* = \varphi\}. \tag{12}$$

Further, by induction, it follows from Eq. ([12](#)) that

$$\Pr\left\{A_n \cap D_\alpha^* = \varphi\right\} \leq (1 - \eta)^{n - N_0}. \tag{13}$$

6 Experimental Study

Our experiments are executed on a personal computer with CPU/3.00 GHz and RBM/512 MB. In order to examine SAIOA's performance characteristics, three well-known mutation rules of polynomial mutation, classical Gaussian mutation and nonuniform mutation are designated in order as its mutation operations; correspondingly, three similar optimization approaches are acquired, where we simply call them SAIOA-P, SAIOA-G and SAIOA-U respectively. Take $N = 40$, $m_0 = 20$, and $m_n = 30$. Also, take the maximal iteration number as 200.

Example Feed Mixer Problem [2]

$$Max\, f(x) = E[24.55x_1 + 26.75x_2 + 39.00x_3 + 40.50x_4 + \eta],$$

$$s.t., \begin{cases} \Pr\{x_1\xi_1 + x_2\xi_2 + x_3\xi_3 + x_4\xi_4 \geq 21\} \geq 1 - \alpha, \\ 2.3x_1 - 5.6x_2 - 11.1x_3 - 1.3x_4 \geq 5, \\ x_1 + x_2 + x_3 + x_4 = 1, x_1, x_2, x_3, x_4 \geq 0, \\ \eta \sim N(0, 2), \xi_1 \sim N(12, 0.53), \xi_2 \sim N(11.9, 0.44), \\ \xi_3 \sim N(41.8, 4.5), \xi_4 \sim N(52.1, 0.79). \end{cases}$$

The difficulty of solving such CCP model consists in the fact that the random factors influence seriously the optimized quality. Such model can be transformed into an analytically equivalent nonlinear optimization one, and thus we can find its

Fig. 1 Comparison of average search curves
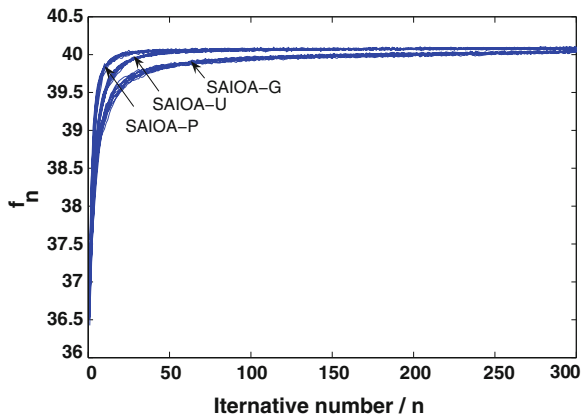
theoretical minimum 39.9337 whatever α takes between 0 and 1. Next, the above three algorithms directly run respectively 50 times on the sample average approximation model P_α^m of the above CCP. They can all obtain reliable solutions during each execution and with any given significance level between 0 and 1. Figure 1 displays the average search curves.

Although the optimized qualities acquired by the approaches are influenced by noises, SAIOA-P and SAIOA-U can find the same optimal reliable solution for any execution and each designated significance level, and appear stable search performances because of extremely small variances caused. Thus, they share abilities of strong exploitation, exploration and noisy suppression. However, SAIOA-G can only find sub-optimal solutions, as Gaussian mutation results in its weak ability of global exploitation. All these demonstrate that SAIOA is a useful optimization framework for chance-constrained programming, and meanwhile Fig. 1 also demonstrates that it is convergent.

7 Conclusions and Further Work

We in this paper give a sufficient condition for existence of solutions of CCP, and also analyze the approximate relation between the true CCP and the sample average approximation problem. We next develop an available and extensible immune optimization approach for CCP problems, while its convergence is proven to be true. Despite of strongly noisy disturbance, experimental results show that SAIOA is convergent with probability 1. It also displays many merits, e.g., simplicity, low computational complexity, and convergence. Comparative experiments illustrate that different mutation rules make SAIOA expose different performance characteristics; SAIOA-A and SAIOA-U perform well over SAIOA-G; relatively, SAIOA-U is a more potential tool for complex CCP problems than SAIOA-P.

References

1. Jin Y, Branke J (2005) Evolutionary optimization in uncertain environments-a survey. IEEE Trans Evol Comput 9(3):303–317
2. Poojari CA, Varghese B (2008) Genetic algorithm based technique for solving chance constrained problems. Eur J Oper Res 185:1128–1154
3. Zhang H, Li P (2011) Chance constrained programming for optimal power flow under uncertainty. IEEE Trans Power Syst 26(4):2417–2424
4. Kampempe JDB, Luhandjula MK (2012) Chance-constrained approaches for multiobjective stochastic linear programming problems. Am J Oper Res 2:519–526
5. Nakama T (2009) A markov chain that models genetic algorithms in noisy environments. Nonlinear Anal 71:991–998

6. Chen CH, He DH (2008) Efficient simulation budget allocation for selecting an optimal subset. J Comput 20(4):579–595
7. Hart E, Timmis J (2008) Application areas of AIS: the past, present and the future. Appl Soft Comput 8(1):191–201
8. Zhang Z (2011) Noisy immune optimization for chance-constrained programming problems. Appl Mech Mater 48–49:740–744
9. Pagnoncelli BK, Ahmed S, Shapiro A (2009) Sample average approximation method for chance constrained programming: theory and application. J. Optim Theor Appl 142:399–416

Robust Coordination with Transmission Delay and Measurement Noises

Dequan Li, Xinyu Huang and Zhixiang Yin

Abstract This paper is concerned with the problem of distributed stochastic approximation in single-integer multi-agent systems on general directed *unbalanced* networks with measurement noises and transmission delay. The time-varying control gains satisfying the stochastic approximation conditions are introduced to attenuate noises. Then based on Lyapunov technique, the convergence result of mean square consensus is established provided that the transmission delay is bounded.

Keywords Consensus · Delay · Stochastic approximation · Measurement noise

1 Introduction

Consensus seeking is of fundamental importance in distributed coordination of multi-agent systems [1, 2]. Therefore, consensus problem has gained a significant interest in the last decade and has been studied under a wide variety of conditions (such as networks with undirected or directed links, switching or fixed topology). Consensus problem involves designing networked interaction protocols that enable agents reaching a global agreement. Being a special case, average consensus is important in many applications and requires agents agreeing on the exact average of their initial states.

D. Li (✉) · X. Huang · Z. Yin
School of Science, Anhui University of Science and Technology,
Huainan 232001, Anhui, People Republic of China
e-mail: leedqseu@gmail.com

Z. Yin
e-mail: zxyin66@163.com

In practical applications, information communication between interacting agents of networks may suffer from transmission delays and measurement noises. Thus, robust consensus protocols that can cope with the joint effects of delays and noises are of more practical interest. However, up to now, most of contributions have just separately dealt with the effect of transmission delays and measurement noises on consensus. It is now known that stochastic approximation with decreasing control gains is a powerful tool to deal with noisy measurements. With which, each agent can gradually reduce the weights assigned to its neighbors and thus attenuate the measurement noises. Based on this idea and by converting the delayed system into a delay less one via the augment method, [3] recently proposed stochastic approximation protocols for randomly switching directed networks in the presence of both communication delay and noisy measurements, whereby convergence of mean square consensus is conducted via ergodic backward products of compatible nonnegative matrices. Stochastic approximation protocols are also adopted for non-leader–follower and leader–follower multi-agent systems on undirected networks [4]. In order to overcome the difficulties induced by the delays and noises, an auxiliary system is introduced via the augment method. Then based on the auxiliary system and by the algebraic theory, the robust consensus problem is transferred into that of analyzing the consensus for a delay-free system with input of color noises. Similarly, via the augment method and using results from non-reversible Markov chains [5], the impact of delays on consensus for directed networks is recently characterized by [6]. Without using the augment method, [7] recently proposed an effect stochastic approximation protocol to deal with transmission noises and bounded delay of multi-agent system on directed balanced networks, where the convergence analysis is mainly based on matrix decomposition. It is noted that, except that [3], many existing distributed stochastic approximation protocols can only be applicable to directed balanced networks, where the corresponding adjacent matrices must be *doubly stochastic*. However, using *doubly-stochastic* matrices implies a feedback communication between a pair of agents which is not always possible and may bring about implementation issue in practice. In this paper, motivated by [7] and [8], we develop a stochastic Lyapunov approach to analyze robust consensus over directed *unbalanced* networks for the case of coexistence of transmission noises and bounded delay, and our aim is also to make a contribution within the stochastic approximation context.

The rest of this paper is structured as follows: In Sect. 2, we formally state the problem of interest and briefly reviews some preliminaries on graph theory. In Sect. 3, we deduce the mean square consensus. Then finally is the concluding remark.

2 Problem Formulation

We consider a multi-agent network with agent (or node) set $\mho = \{1, 2 \ldots, N\}$, for which each agent has the following discrete-time single-integrator dynamics

$$x_i(k+1) = x_i(k) + u_i(k), \quad k = 0, 1, 2, \cdots, \quad i = 1, 2, \cdots, N \qquad (1)$$

where $x_i(k)$ is the state of agent i, $u_i(k)$ is the control input or protocol of agent i.

The topology of the communication network can be modeled by a directed graph (or digraph) $\mathbf{G} = (\mho, \mathcal{E}, W)$, where $\mathcal{E} = \{e_{ij} = (i,j)|i, j \in \mho\}$ denotes the edge set and the corresponding weight adjacent matrix is $W = (w_{ij}) \in R^{N \times N}$. The edge $e_{ij} \in \mathcal{E}$ represents that agent i can receive information from agent j, then agent j is called an in-neighbor of agent i and $w_{ji} > 0$, otherwise $w_{ji} = 0$. A directed graph is *undirected* if $e_{ij} \in \mathcal{E} \Leftrightarrow e_{ji} \in \mathcal{E}$ for $\forall i \neq j$. A digraph $\mathbf{G}$ is strongly connected if for any distinct nodes i and j, there exists a path that connects i and j. A digraph $\mathbf{G}$ is balanced if the in-degree and out-degree of each $i \in \mho$ are same. W is said to be nonnegative if its entry $w_{ij} \geq 0$ for all i and j. W is said to be stochastic if it is nonnegative and satisfies $W\mathbf{1} = \mathbf{1}$, where $\mathbf{1} = (1, 1, \cdots, 1)^T \in R^N$. Moreover, W is called doubly stochastic if it is stochastic and satisfies $\mathbf{1}^T W = \mathbf{1}^T$.

Due to the random communication environment and time delay, at each time, agent i receives the delayed noisy measurements of its neighbor j's state given by

$$y_{ji}(k-\tau) = x_j(k-\tau) + \xi_{ji}(k-\tau) \qquad (2)$$

where $\xi_{ji}(k-\tau)$ describes the measurement noise affecting the information transmission along the directed link e_{ji} at time k. $x_j(k)$ is the state of agent j; τ is the time delay.

Then the control input for agent i at time k is proposed as

$$u_i(k) = \alpha(k) \sum_{j=1, j \neq i}^{N} w_{ij}\left(y_{ji}(k-\tau) - x_i(k-\tau)\right) \qquad (3)$$

which means that agent i updates its current state by taking a weighted average of its own delayed state and the delayed noisy states of its neighbors. Where w_{ij} are the entries of the stochastic adjacent matrix W associated with the digraph $\mathbf{G}$. The scalars $a(k) \in (0, 1](k = 0, 1, \cdots)$ are the time-varying control gains, which are used to attenuate measurement noises.

Before further proceeding, we need the following assumptions.

Assumption 1 In the stochastic approximation approach, the control gain sequence $\{\alpha(k), k \geq 0\}$ satisfies the following:

$$\sum_{k=0}^{\infty} \alpha(k) = \infty \text{ and } \sum_{k=0}^{\infty} \alpha^2(k) < \infty$$

Assumption 2 The digraph G is strongly connected. W is stochastic and has positive diagonal entries, that is, there exists a positive constant ρ such that $w_{ii} = 1 - \sum_{j=1,\,j\neq i}^{N} w_{ij} \geq \rho > 0$ holds for each $i \in \mho$. Meanwhile, $w_{ij} \in \{0\} \cup [\rho, 1]$ for $i, j \in \mho$ with $i \neq j$.

Assumption 3 There exists a positive integer d^* such that $0 \leq \tau \leq d^*$.

Assumption 4 All the (delayed) measurement noises are zero mean and have uniformly bounded variance, that is, $E\big[\xi_{ji}(k - \tau)\big] = 0$ and $E\big[\xi_{ji}^2(k - \tau)\big] \leq \sigma^2$ for all $i, j \in \mho$, $k \geq 0$ and $0 \leq \tau \leq d^*$, where $E[\cdot]$ denotes the expectation operator. Moreover, $E\big[\xi_{ji}(k-\tau)\xi_{l\kappa}(t-\tau)\big] = 0$ for all $i, j, l, \kappa \in \mho$ and $k \neq t$.

Applying the protocol (2) to (1) and by Assumption 2, we obtain the closed-loop system for agent i:

$$x_i(k + 1) = x_i(k) + \alpha_k \sum_{j=1,\,j\neq i}^{n} w_{ij}x_j(k - \tau) - \alpha_k \sum_{j=1,\,j\neq i}^{n} w_{ij}x_i(k - \tau) + \alpha(k)\xi_i(k-\tau)$$

$$= x_i(k) - \alpha_k x_i(k - \tau) + \alpha_k \sum_{j=1}^{n} w_{ij}x_j(k - \tau) + \alpha(k)\xi_i(k-\tau)$$

$$(4)$$

where $\xi_i(k-\tau) = \sum_{j=1,\,j\neq i}^{N} w_{ij}\xi_{ji}(k-\tau)$ denotes the aggregated delayed measurement noises that agent i receives from all of its neighbors at time k.

Denote $x(k) = (x_1(k), \cdots, x_N(k))^T$, $\xi(k) = (\xi_1(k), \cdots, \xi_N(k))^T$. Then (4) can be further written in a compact form

$$x(k + 1) = x(k) + \alpha(k)(W - I)x(k-\tau) + \alpha(k)\xi(k - \tau) \qquad (5)$$

the initial states are $x(0), x(1), \cdots, x(\tau)$ and $\xi(l) = 0$ for $l < 0$. Furthermore, denote the regression function $H(x(k)) = (W - I)x(k)$, then it is yielded from (5) that

$$x(k + 1) = x(k) + \alpha(k)[H(x(k - \tau)) + \xi(k - \tau)] \qquad (6)$$

Thus, by Assumption 4, we can easily obtain the following results for the aggregated delayed noise $\xi_i(k - \tau)$, that is, $E[\xi_i(k - \tau)] = 0$ and $E\big[\xi^T(k - \tau)\xi(t - \tau)\big] = 0$ for $k \neq t$. Meanwhile, by the definition of aggregated delayed noise $\xi_i(k - \tau)$ and Assumptions 2 and 4, there holds that

$$\big|E\big[\xi_i(k - \tau)\xi_j(k - \tau)\big]\big| \leq \sigma^2(1 - w_{ii})(1 - w_{jj}) \leq \sigma^2(1 - \rho)^2 \qquad (7)$$

Remark 1 Note that W is stochastic and has positive diagonal elements such that G is strongly connected. Then by [9], W is primitive and 1 is a simple maximal eigenvalue of W with algebraic multiplicity one. Furthermore, there exists unique

normalized positive left eigenvector $\pi = (\pi_1, \cdots, \pi_N)^T$ corresponding to eigenvalue 1 and satisfies

$$\pi^T W = \pi^T, \pi^T 1 = \sum_{i=1}^{N} \pi_i = 1 \text{ and } \lim_{k \to \infty} W^k = 1\pi^T \tag{8}$$

If W is doubly stochastic, then $\pi^T = \frac{1}{N}(1, 1, \cdots, 1)^T$.

For analyzing the convergence of (5), we introduce the following definition.

Definition 1 (*Mean Square Consensus* [3, 4]). All agents are said to reach mean square consensus if there exists a common random variable x^* such that $\lim_{k \to \infty} E|x_i(k) - x^*|^2 = 0$ and $E\|x(k)\|^2 < \infty$ for all $i \in \mho$ and $k \geq 0$.

3 Convergence Analysis

Before going further, we need some useful lemmas that will be used in the proof of our result.

Lemma 1 [7] *Suppose that Assumptions 1, 2 and 3 hold, then* $\sum_{k=\tau}^{\infty} \alpha(k)\xi(k-\tau) < \infty$ *and* $H(x(k-\tau)) - H(x(k)) \to 0$ *as* $k \to 0$.

Lemma 2 [10] *Suppose that Assumption 1 holds, meanwhile*

(a) $\chi(k) = \varpi(k) + \sigma(k)$, where $\varpi(k) \to 0$ as $k \to 0$ and $\sum_{k=0}^{\infty} \alpha(k)\sigma(k) < \infty$, $\alpha(k)$ is defined in Assumption 1;

(b) The matrix B only has negative and zero eigenvalues with the zero eigenvalue having the same algebraic multiplicity and geometric multiplicity.

Then the sequence $y(k+1) = y(k) + \alpha(k)(By(k) + b + \varpi(k) + \sigma(k))$ satisfies that $\lim_{k \to \infty} \text{dist}(y(k), \Xi) = 0$, where $\Xi = \{y|By + b = 0\}$ and $\text{dist}(y(k), X) = \inf_{x \in X} \|y(k) - x\|$.

Remark 2 It yields from Remark 1 that the matrix $W - I$ only has negative and zero eigenvalues with the zero eigenvalue having the same algebraic multiplicity and geometric multiplicity.

Note that (5) can be further rewritten as

$$\begin{aligned} x(k+1) &= x(k) + \alpha(k)(W-I)x(k) + \alpha(k)[H(x(k-\tau)) - H(x(k))] + \alpha(k)\xi(k-\tau) \\ &= P(k)x(k) + \alpha(k)[H(x(k-\tau)) - H(x(k))] + \alpha(k)\xi(k-\tau) \end{aligned} \tag{9}$$

where $P(k) = (1 - \alpha(k))I + \alpha(k)W$. Note the fact that W is stochastic and $a(k) \in (0, 1](k = 0, 1, \cdots)$ are time-varying, and thus $P(k)$ is also stochastic and time-varying; meanwhile, $P(k)$ also satisfies (8).

Denote the following Lyapunov function

$$
\begin{aligned}
V(x(k)) &= x^T(k)\big(I - \pi 1^T\big)D\big(I - 1\pi^T\big)x(k) = x^T(k)\big(D - \pi\pi^T\big)x(k) \\
&= \sum_{i=1}^{N} \pi_i\big(x_i(k) - \pi^T x(k)\big)^2
\end{aligned}
\tag{10}
$$

where the diagonal matrix $D = diag(\; \pi_1 \quad \cdots \quad \pi_N\;)$. The second equality in (9) is yielded from $1^T D = \pi^T$, $D1 = \pi$ and $\pi^T 1 = 1$. The function $V(x(k))$ is actually a weighted distance Lyapunov function that measuring the weighted spread of the vector $x(k)$ components with respect to the weighted average value $\pi^T x(k)$.

Lemma 2 [8] *If Assumption 2 holds, then*

$$
\begin{aligned}
V(x(k+1)) &= x^T(k)P^T(k)\big(D - \pi\pi^T\big)P(k)x(k) \\
&\leq \left(1 - \frac{\eta(k)}{2(N-1)}\right)V(x(k))
\end{aligned}
\tag{11}
$$

where $1 > \eta(k) = \pi_{\min}\rho\alpha(k) > 0$ and $\pi_{\min} = \min_{i \in \mho} \pi_i$.

With these lemmas in place, then under the proposed protocol (2), all agents of the directed network will reach mean square consensus.

Theorem 1 *If Assumptions 1-4 hold, then for each agent i $(i \in \mho)$, there holds*

$$
\lim_{k \to \infty} E|x_i(k) - x^*|^2 = 0
$$

with $E(x^*) = \hat{\pi}^T \hat{x}(\tau)$, where $\hat{\pi}$ satisfies $\hat{\pi}^T P = \hat{\pi}^T$. That is, all the agents converge in mean square to a common random variable x^*, and its mathematical expectation is the weighted average of agents' initial states; furthermore, the variance of x^* is bounded.

4 Conclusion

The paper considers the robust consensus algorithm that performs correctly despite coexistence of transmission noises and bounded delay. We do not model the behavior of the bounded delay using any virtual node. For future work, we will characterize the relationship between the convergence rate and the time-varying control gains.

References

1. Jadbabaie A, Lin J, Morse AS (2003) Coordination of groups of mobile autonomous agents using nearest neighbor rules. IEEE Trans Autom Control 988–1001
2. Ren W, Beard RW (2005) Consensus seeking in multi-agent systems under dynamically changing interaction topologies. IEEE Trans Autom Control 655–661
3. Huang M (2012) Stochastic approximation for consensus: a new approach via ergodic backward products. IEEE Trans Autom Control 2994–3008
4. Liu S, Xie L, Zhang H (2011) Distributed consensus for multi-agent systems with delays and noises in transmission channels. Automatica 920–934
5. Fill JA (1991) Eigenvalue bounds on convergence to stationarity for non reversible markov chains, with an application to the exclusion process. Ann Appl Probab 62–87
6. Tsianos KI, Rabbat MG (2011) Distributed consensus and optimization under communication delays. In: Proceedings of 49th Allerton conference on communication, control and computing, pp 974–982
7. Xu J, Zhang H, Shi L (2012) Consensus and convergence rate analysis for multi-agent systems with time delay. In: Proceedings of 12th international conference on control, automation, robotics and vision, pp 590–595
8. Li DQ Wang XF (2013) Robust consensus for multi-agent systems over unbalanced directed networks. J Syst Sci Complex
9. Horn RA, Johnson CR (1985) Matrix Analysis. Cambridge University Press, Cambridge
10. Zhu YM (1985) Stochastic approximation under a class of measurement noises. Acta Mathematica Scientia 5:87–98

Study on the Importance of Cultural Context Analysis in Machine Translation

Zonghua Wang

Abstract Context of culture affects the specific meaning of the language. So the analysis of cultural context is essential for Machine Translation (MT). If the cultural context analysis of the source language is omitted in MT, ambiguity or mistranslation will be produced. The paper analyzes briefly the present situation of MT first and then, based on the theory of cultural context, it discusses with examples the importance of cultural context analysis in MT. The paper holds that if the cultural context analysis is paid enough attention in MT, the literal meaning and cultural connotation of source language can be exactly and completely translated into target language so as to improve the quality of MT. Therefore, the large-scale context corpus is very necessary for MT.

Keywords Machine translation · Cultural context analysis · Context corpus

Machine Translation (MT) is not only the conversion between natural languages, but the communication between different cultures. Because language is a part of culture while culture is the soil for language to grow up gradually. Context of culture affects the specific meaning of the language. So the analysis of cultural context is essential for MT. If the cultural context analysis of source language is omitted in MT, ambiguity or mistranslation will be produced. The paper discusses with examples the importance of cultural context analysis in MT. The paper holds that if the cultural context analysis is paid enough attention in MT, the literal

Z. Wang (✉)
School of Foreign Languages, Anhui University of Science and Technology,
Huainan, People's Republic of China
e-mail: Wanggmp6@yahoo.com.cn

Z. Wang
School of History and Culture, Shandong University, Jinan, People's Republic of China

meaning and cultural connotation of source language can be exactly and completely translated into target language so as to improve the quality of MT. Therefore, the large-scale context corpus is completely necessary for MT.

1 Machine Translation: A Brief Introduction

Machine Translation (MT) is the translation by computer between natural languages including their texts and sounds, which is the time-honored section of natural language processing by computer and "the most important application field in Computational Linguistics" [1]. Besides, MT is a comprehensive discipline combined with linguistics, mathematics and computer science in the information age. The technique of language analysis is the core of MT [2]. At present, MT system can be classified into Rule-based Machine Translation system, Corpus-based Machine Translation system, Multi-engine Machine Translation system and Spoken Language Machine Translation system, among which Rule-based Machine Translation system is classified into Literal Translation system, Transition system and Interlingua system, and Corpus-based Machine Translation system into Example-based Machine Translation system and Statistical Machine Translation system [3].

In the early 1930s, French engineer G. B. Artsouni put forward the idea of translation by machine and got a patent for "Translation Machine" named "Mechanical Brain". In 1946, the first Electronic Numerical Integrator and Computer (ENIAC) in the world came out. Its fantastic speed at that time indicated the innovation of translation technique. In March, 1947, Warren Weaver, an American Scientist, and Andrew Booth, a British engineer, met together at Rockefeller Center in New York to discuss the computer application problem. In 1949, Warren Weaver issued a memorandum titled *Translation,* in which he presented formally the problem of computer translation. And then some scholars from USA, Soviet Union, Britain and so on took part in the research on machine translation in following several years. Georgetown University and IBM jointly carried out the experiment of machine translation first in the world in 1954 and succeeded in translating several simple Russian sentences into English, which seemingly indicated that the high-qualified MT software would be made very soon and there was no language barrier in the future.

China is one of the countries earlier carrying out the study of MT, beginning with its research work in 1956. The study of MT approximately goes through four periods: the beginning period (from the middle 1950s to the middle 1960s), the stagnant period (from the middle 1960s to the middle 1970s), the developing period (from the middle 1970s to the late 1980s) and the flourishing period (from the early 1990s to the present), which has made some achievements in theory and practice. However, there is a long way for Chinese experts to go to catch up with and even surpass the international level. One reason is that Chinese is a more complicated language than English, one of Indo-European languages, making it

more difficult in MT studies. At present, many experts have realized that the study of Chinese itself is basic for MT from Chinese to other languages and vise versa.

It is almost 60 years since the first MT experiment was successfully made in 1954. In the past 60 years, MT made great progress in many countries. With the development of information technology and the popularization of internet and computer, MT application has been spread increasingly, such as Canadian TAUM-METEO (which can automatically receive, count and translate the data of weather forecast, and finally issue it in English and French.), automatic translation telephone made by UK, USA, Germany, etc., electronic dictionary, net-automatic translation system, GPS for car drivers (based on speech recognition, transformation and synthetic system) and so on. Now anyone can use the MT tools on the internet. MT plays an important role instead of human in the translation of sci-tech, political documents and so on, and increases efficiency greatly.

However, people are not satisfactory with the quality of MT at the present time. Reasons for this are the complexity of natural languages and their translation and the limitations of computer, e.g., smaller storage capacity and less complicated than the human brain [4]. Nida, a famous American translation theorist, points out that in addition to problems of size and sophistication, there is a need of more appropriate intermediate technical language which can be automatically adapted to the special limitations of computer and has more comprehensive function than the present machine languages, e.g. Comit, Fortran and Mimic [5]. Among the reasons, the failure of cultural context analysis in MT is a very important factor for producing low-qualified versions.

2 The Theory of Context of Culture

Context of culture is one of the most important linguistic categories in the study of language use and function, which is first proposed by British anthropologist Bronislaw Malinowski. It refers to the specific social regulations and idioms of certain speech community. The recent research of cultural context shows that each speech community has its own history, culture, customs, social regulations, mode of thinking, moral ideas and value orientation. These specific ways and factors of certain speech community constitute the full content of cultural context [6].

Language is a very important part and the carrier of a culture, which is indispensable for both the functioning and the perpetuation of the culture; culture is the surroundings for language to grow up and determine its specific linguistic meaning. "Language also constitutes the most distinctive feature of a culture, which may be described in a simplistic manner as the totality of the beliefs and practices of a society" [7]. They are closely connected and indivisible. Juri Lotman said, "No language can exist unless it is steeped in the context of culture" [8]. Words only have meaning in terms of the corresponding culture. Susan Bassnett also said, "The translator treats the text in isolation from the culture at his peril" [8]. For example, he was in the seventh heaven last night. If the sentence is literally translated into Chinese as "他昨天晚上在第七天堂", no Chinese reader

can understand the version because the phrase "the seventh heaven" does not mean itself literally but refers to the God's Palace, meaning here "the state of great happiness", and it means nothing practically in Chinese. Another Example, at the banquet, Mr. Li jokingly said to a foreign guest, pointing at Mr. Zhang next to him, "他有气管炎". All Chinese guests burst into laughing. The young interpreter translated it literally as follows: "He suffers from tracheitis". The foreign guest was in a fog and did not laugh, and asked confusedly, "Why are you laughing so hard when he is suffering from tracheitis?" Obviously, the interpreter makes a pragmatic mistake without considering Chinese cultural context. "气管炎" and "tracheitis" mean the same. But "妻管严" and "气管炎" are homophones of each other in Chinese. "妻管严" with Chinese cultural connotation means hen-pecked.

It is thus clear that "translation is the communication between two cultures. For the successful translation, to know two cultures well is more important than to know their languages" [9]. The meaning of a text is finally determined by its cultural context, not by its external language form. Therefore, while translating, the translator should take the cultural context of source language into consideration in order to understand fully the source text and translate it exactly into the target language; at the same time, readers' cultural context should be considered before final version is produced so that they can understand the target text fully [10]. It is also true for machine translation.

3 Cultural Context Analysis and Machine Translation

It is easily seen from the analysis mentioned above that context of culture plays a very important role in translation. Similarly, MT can not be isolated from the cultural context analysis. Otherwise, the source language text can not be understood exactly and fully and mistranslation will be produced in MT. That is to say, the cultural context analysis is the extremely important condition in MT for computer to make a choice of the given versions. It affects MT's quality and even determines whether the version is right or wrong. The following are four examples (two in English and two in Chinese) to show us the importance of study on cultural context analysis and its application in MT. Versions come from English to Chinese and Chinese to English MT soft systems such as Google online translation system, Baidu online translation system and Kingsoft Fast system which are very popular and easily available in China.

Example 1: He is by no means courageous. He is actually a hare.

Google Version: 他绝不是勇敢的。他其实是一只野兔。
Baidu Version: 他绝不是勇敢的。他实际上是一个兔子。
Kingsoft Version: 他绝不是勇敢的。他实际上是一个兔子。

The word "hare" in English refers figuratively to a person who is very timid. However, the word "mouse" is used in the same way instead of "hare" in Chinese. Three Chinese visions from MT systems are similar. They are all literal translations with English cultural connotation missed and do not express exactly the

meaning of the source text. If the word "hare" is translated into "胆小如兔" instead of "一只兔子", Chinese readers can not understand the version either because Chinese cultural connotation is missed here. So reference translation is "他绝不是勇敢的。他其实胆小如鼠"。

Example 2: Our history teacher is a good talker. I'd like to ask him, "where is the beef?" because we sure don't learn much from him.

Google Version: 我们的历史老师是一个很健谈。我问他，"是牛肉在哪里?"因为我们肯定不会学到很多东西从他身上。
Baidu Version: 我们的历史老师是个很健谈的人。我想问他，"牛肉在哪里?"因为我们当然不从他的身上学到很多。
Kingsoft Version: 我们的历史老师是个很健谈的人。我想问他，"牛肉在哪里?"因为我们当然不从他的身上学到很多。

The short sentence "where is the beef?" in source text is an American idiom, meaning that people complain there is much too little beef in hamburger. Now it refers to something that is short of substantiality. Computer does not know of the cultural background of the idiom. As a result, versions in the Example 2 apparently shows us the literal meaning of the idiom while neglecting its cultural connotation and that there is no the cultural context analysis of source text in three MT systems. Therefore, Chinese versions are very hard for readers to understand. The following version is better with the cultural context analysis of both source text and target text:

"我们的历史老师很能夸夸其谈。我真想问问他，"管用的货色在哪里?因为我们真的没有从他那里学到多少东西" [11]。

Example 3: 三个臭皮匠抵个诸葛亮。

Google Version: Three Stooges arrived Zhuge Liang.
Baidu Version: The Three Stooges against Zhu Geliang.
Kingsoft Version: The Three Stooges against Zhu Geliang.

In Chinese culture, "诸葛亮" (Zhuge Liang) is a person of great wisdom and resourcefulness; "皮匠" (cobbler or shoemaker) refer to less educated people who make a living on repairing shoes. Sometimes,people call sneeringly such less educated people as "臭皮匠". The cultural connotation of source text is "more people, more wisdom". All three MT versions are unacceptable because they all contain no cultural connotation of source text. Second and third MT systems even wrongly translate "诸葛亮", whose family name is "Zhuge", not "Zhu". By analyzing the cultural context of source text, the acceptable version is "Three cobblers with their wits combined equal Zhuge Liang, the mastermind". At the same time, there is the same expression in English as "Two heads are better than one". Considering the cultural context of target language, this is better version for target text readers.

Example 4: 他们俩真是一龙一猪!

Google Version: The two of them is really a dragon, a pig.

Baidu Version: Both of them are Yilongyizhu!
Kingsoft Version: Both of them are Yilongyizhu!

Both "龙" (dragon) and "猪" (pig) in English are derogatory terms while the word "龙" (dragon) in Chinese is commendatory term. The set phrase "一龙一猪" in Chinese figuratively refers to two persons, one of whom is bright, capable and promising, and the other is extremely foolish and hopeless. "一龙一猪" is literally translated in Google version and it is just expressed again in another Chinese form —Pinyin in both Baidu and Kingsoft versions, not translated. With the cultural context lost, three MT versions do not express the cultural connation of source text and make it very difficult for target text readers to know what is really said. The similar or same English expression as "一龙一猪" is not available as well. Maybe, the following version is acceptable: "They are sharp contrastive individuals: one is very capable, while the other is extremely incompetent".

4 Conclusion

In conclusion, cultural context analysis plays an extremely important role in MT. The main reason for the low quality of MT in the examples mentioned above is the failure of cultural context analysis of source language. A computer does not know of the cultural differences between languages and naturally it can not understand the context of source text. Therefore, the large-scale bilingual or multilingual context corpus is very necessary for MT to improve its version quality. Just as Newmark said, "context is an overriding factor in all translation, and has primacy over any rule, theory or primary meaning" [12]. So context corpus should contain an extremely huge number of versions related to the context of utterance and culture as basic language sources for computer. With such context corpus, computer can "analyze" like a human translator such contexts of source text, compare the different versions input and choose the closest correspondence with the source text. However, to set up such context corpus will be an enormous system engineering. With the further development of informationalization and popularization of network in the 21st century, people need more information by computer from different nations, which will certainly promote the research on MT and the really full-automatic and high qualified MT soft system will surely be generated in the future so as to overcome the barrier of languages in the cross-cultural communication for information.

References

1. Zhiwei F (2011) Computational linguistics: its past and present. J Foreign Lang 1:11
2. Shanghui K, Sumin Q (2012) An investigation on machine translation. Intell Comput Appl 2 (42)

  3. Zhiwei F (2004) Machine translation studies. China Translation and Publishing Corporation, Beijing
  4. Zheng Z (2005) Difficulties for machine translation. Foreign Lang Res 5:59–62
  5. Nida EA (2004) Toward a science of translation. Shanghai Foreign Language Education Press, Shanghai
  6. Lan H (2008) Cultural context: the hedge of version choice. J Guangxi Soc Sci 6:175
  7. Nida EA (2001) Language and culture: contexts in translating. Shanghai Foreign Language Education Press, Shanghai
  8. Bassnett S (2004) Translation studies. Shanghai Foreign Language Education Press, Shanghai
  9. Liu W (2012) On cultural context factors affecting translation. J Liaoning Univ Technol (Soc Sci Ed) 1(52)
 10. Wang J (2006) Cultural context and translation. Shanghai J Translator 2:52–54
 11. Xueyun E, Chen X (2007) Context analysis: the indispensable factor in machine translation. Foreign Lang Teach 4(53)
 12. Newmark P (2001) Approaches to translation. Shanghai Foreign Language Education Press, Shanghai

The Evaluation Model of Construction Industry Financial Risk Based on SVM

Xiaomei Chang

Abstract The model of evaluating profession financial risk can be added with industry characteristic indicator, a general evaluation model is more suitable for a particular industry, and it can improve the accuracy. This paper has chosen 23 financial indexes aims at six respects such as solvency, profitability, cash flow, asset management ability, development ability and capital structure of construction industry, and has built the evaluation index system of building trades. Selected the construction enterprises in the Shanghai and Shenzhen listed as samples, applied cluster analysis on samples to preprocessing, then using the theory of support vector machine, respectively constructed polynomial Gauss radical, Sigmoid three kernel function SVM evaluation model of construction industry financial risk, and has verified them, The results shows that the SVM model that applied in the financial risk evaluation of construction industry has a high accuracy.

Keywords Construction industry · Financial · Cluster analysis · Support vector machine · Kernel function

1 Introduction

In recent years, the research which on enterprise financial risk model mainly concentrated on the general model widely used are statistical model and artificial intelligent model. Statistical model contains Z score model, multivariate discriminate analysis model, Logistic regression model and principal component analysis model. Artificial intelligence model mainly contains fuzzy evaluation model and neural network model [1].

X. Chang (✉)
School of Economics and Technology, Anhui University of Science and Technology, Huainan, 232001 Anhui, China
e-mail: 573748396@qq.com

Management model, organization model, products and risk of construction industry have characteristics different from other industry have characteristics different from other industry. Using a general model to evaluate the financial risk will be biased. Constructing a financial risk evaluation model that specialized for construction industry to monitor and evaluate it's financial situation, it has important significance for guarding against and defusing financial risk of construction enterprises [2]. But the above model in application must depend on a lot of Historical samples, it will affect its accuracy for construction industry which has less data.

This paper aimed at the realistic situation of less samples of construction enterprise, using small sample learning method-support vector machine (SVM-Support Vector Machines), has established the evaluation model of construction industry financial risk based on SVM, in order to improve the accuracy of evaluation.

2 The Selection of Sample and Cluster Analysis

2.1 The Selection of Sample

The study samples are A shares of construction and whose enterprises that listed on the Shanghai and Shenzhen stock exchange. This paper have selected 38 companies that are listed in Shanghai and Shenzhen, whose financial data have announced and whose main businesses are civil, design planning, engineering contracting and real estate development. The datum in samples are directly obtained from the public annual accounting report or obtained by processing.

2.2 Cluster Analysis

2.2.1 The Theory of Cluster Analysis

Cluster is composed data object into different classes, so that the similarity between different types become smaller while similarity between similar object become larger [3]. Cluster analysis is a kind of common statistical analysis method that is studying (sample or index) the classification problems. Clustering is an unsupervised, observed learning process, it not gen people's subjective judgment and based on data, so it is objective and fair, cluster analysis can be separately as a tool to obtain the distribution of data, then through analyze the data, it also can be used as a preprocessing step for other algorithms. In this study, before building evaluation model of construction enterprise financial risk, using cluster analysis to

classify the selected 38 construction enterprises without a priori, so that avoid classifying the samples confine within ST and Non ST.

2.2.2 Application of Cluster Analysis

According to the above evaluation index system and sample data, cluster analyzing relevant financial data of construction enterprises by common hierarchical cluster method.

Through the cluster analysis of spss17.0, the risk finance of construction not only covers the construction enterprise that has been defined as ST, such as ST Huitong, but also Not ST construction enterprise, contains: Shenzhen Universe Group Co. Ltd, NORINCO International, Xiamen Academy of Building Research Group Co., Ltd, Shanghai Pudong Road and Bridge Construction Co. Ltd, Longyuan Construction Group Co. Ltd, Tengda Construction Group Co. Ltd, Longjian Road and Bridge Co. Ltd, Keda-Group Company, China Railway Construction Corporation Limited. This reflects that using cluster analysis can eliminate the priori interference. Classifying potential financial crisis enterprises into dominant financial crisis enterprises is more predictable.

Now, we constructing a more accurate financial risk evaluation model by using support vector machine based on cluster analysis.

3 The SVM Model of Financial Risk Evaluation

3.1 Support Vector Machine Theory

SVM is a kind of universal learning machine, it is a statistical learning method that realized by Vapnik based on statistical theory, then formed a small samples learning method that based on statistical theory [4].

SVM has a solid theoretical foundation in Mathematics. It is specifically for the small sample learning problems, the method is based on the VC dimension theory of statistical learning theory and structural risk minimization principle. According to the limited information, finding the best compromise between the model complexity (that is the learning accuracy of specific training samples) and the learning ability (that is the capacity to identify any sample without error), it has a stronger generalization ability [5].

The basic idea of SVM is summed up: changing the input space into a high dimensional space by nonlinearity, solving convex two quadratic programming problems with a constraint in this new space, then can obtain the global unique optimal solution.

SVM has the following characteristics:

Model training needs fewer parameter to indentify, and determining the model parameter is an optimization problem of convex objective function.

SVM is a machine learning method of hard structural risk minimization, it finds the structural risk minimization through the generalization error bound, so it has avoided the defect of empirical risk minimization of the general statistical learning methods.

SVM can transforms linear inseparable vector quantity into linear separable one by using the method of mapping low dimensional output space to high dimensional feature space, then calculating the best graphic classification.

SVM using the kernel function of original space, cleverly replaces the inner product operation in the high dimensional feature space, that is $Q_{ij} = \phi(x_i)^T \phi(x_j) = K(x_i, x_j)$, so it can avoid the rapid growing problem in the process of mapping low dimensional output space to high dimensional feature space.

3.2 Establishing SVM Model of Construction Industry Financial Risk Evaluation

For the sample set [6]:

$$(x_1, y_1)\ldots(x_i, y_i)\ldots(x_l, y_l) \in R^N \times R \tag{1}$$

The required function form of fitting:

$$f(x) = W^T \phi(x) + b, \quad W, \phi(x) \in R^N, b \in R \tag{2}$$

where, W is parameter column vector, $\phi(\cdot)$ is column vector of a function, it has mapped input sample to feature space from input space.

The obtained fitting function $f(\cdot)$ need to minimize the following performance index. That structural risk:

$$R_{reg}^{\varepsilon} = \frac{1}{2}\|W\|^2 + C \cdot R_{emp}^{\varepsilon} \tag{3}$$

where, constant C is penalty factor, and $C > 0$, that is the punishment degree of the samples of beyond error, it can get a compromise between training error and the model complexity, so that the obtained function $f(x)$ has a better generalization ability.

Empirical risk:

$$R_{emp}^{\varepsilon} = \frac{1}{l}\sum_{i=1}^{l} |y_i - f(x_i)|_{\varepsilon} \tag{4}$$

Defining the insensitive loss function ε:

$$|y - f(x)|_\varepsilon = \begin{cases} 0, & |y - f(x)| \leq \varepsilon \\ |y - f(x)| - \varepsilon, & |y - f(x)| > \varepsilon \end{cases} \tag{5}$$

Transforming approximation problem of function into the optimization problem as follows:

$$\min \left(\frac{1}{2} W^T W + C \left(v\varepsilon + \frac{1}{l} \sum_{i=1}^{l} \left(\zeta_i + \zeta_i^* \right) \right) \right) \tag{6}$$

$$W^T \phi(x_i) + b - y_i \leq \varepsilon + \zeta_i \tag{7}$$

$$y_i - W^T \phi(x_i) - b \leq \varepsilon + \zeta_i^* \tag{8}$$

$$\zeta_i, \zeta_i^* \geq 0, \varepsilon \geq 0, \quad i = 1, \dots, l \tag{9}$$

where, $0 \leq v \leq 1$, insensitive loss function ε shows that $W^T \phi(x)$ in $y \pm \varepsilon$, and do not account for loss.

The presence of insensitive loss function has increased the model tolerance, compromising maximum of classification interval and minimum of misclassified samples, getting the generalized optimal classification face, obtained the generalized optimal classification face.

Solving the optimization problems, defining Lagrange function:

$$\begin{aligned} L(W, b, \zeta, \zeta^*, \alpha, \alpha^*, \eta, \eta^*) = & \frac{1}{2} \|W\|^2 + Cv\varepsilon \\ & + \frac{C}{l} \sum_{i=1}^{l} \left(\zeta_i + \zeta_i^* \right) - \sum_{i=1}^{l} \left(\eta_i \zeta_i + \eta_i^* \zeta_i^* \right) \\ & - \sum_{i=1}^{l} \alpha_i \left(\varepsilon + \zeta_i + y_i - W^T \phi(x_i) - b \right) \\ & - \sum_{i=1}^{l} \alpha_i^* \left(\varepsilon + \zeta_i^* - y_i + W^T \phi(x_i) + b \right) - \beta\varepsilon \end{aligned} \tag{10}$$

Because $\frac{\partial L}{\partial W} = 0, \frac{\partial L}{\partial b} = 0, \frac{\partial L}{\partial \varepsilon} = 0, \frac{\partial L}{\partial \zeta_i^*} = 0$, so

$$W = \sum_{i=1}^{l} \left(\alpha_i - \alpha_i^* \right) \phi(x_i) \tag{11}$$

$$\sum_{i=1}^{l} \left(\alpha_i - \alpha_i^* \right) = 0 \tag{12}$$

$$Cv - \sum_{i=1}^{l} \left(\alpha_i - \alpha_i^* \right) - \beta = 0, \tag{13}$$

Introducing a kernel function to find the dot product of the feature space $K(x, y) = \phi(x) \cdot \phi(y)$, so obtaining an optimization problem of quadric form:

$$\min\left(\frac{1}{2}\sum_{i,j=1}^{l}\left(\alpha_i^* - \alpha_i\right)K\left(x_i, x_j\right)\left(\alpha_j^* - \alpha_j\right) - \sum_{i=1}^{l}y_i\left(\alpha_i^* - \alpha_i\right)\right) \tag{14}$$

$$st: \sum_{i=1}^{l}\left(\alpha_i^* - \alpha_i\right) = 0 \tag{15}$$

$$0 \le \sum_{i=1}^{l}\left(\alpha_i + \alpha_i^*\right) \le Cv \tag{16}$$

$$0 \le \alpha_i, \alpha_i^* \le C/l, \quad i = 1, \ldots, l \tag{17}$$

We expresses standard quadric form planning type by matrix:

$$\min\left(\frac{1}{2}(\alpha - \alpha^*)^T Q(\alpha - \alpha^*) + y^T(\alpha - \alpha^*)\right) \tag{18}$$

$$st: e^T(\alpha - \alpha^*) = 0, e^T(\alpha + \alpha^*) \le Cv \tag{19}$$

$$0 \le \alpha_i, \alpha_i^* \le C/l, \quad i = 1, \ldots, l \tag{20}$$

where, e is a unit matrix.

Q is a kernel matrix, the elements:

$$Q_{ij} = \phi(x_i)^T \phi(x_j) = K(x_i, x_j) \tag{21}$$

The common kernel function contains polynomial kernel function, Gauss radial basis function, the Sigmoid kernel function and custom function [7].

Polynomial kernel function:

$$K(x, y) = ((x \cdot y) + 1)^d, \quad d = 1, \ldots n \tag{22}$$

Gauss radial basis function:

$$K(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right) \tag{23}$$

where, σ is an undetermined constant.

Sigmoid kernel function:

$$K(x, y) = \tan(\phi(x \cdot y) + \theta) \tag{24}$$

SVM output function:

$$f(x) = \sum_{i=1}^{l} \left(\alpha_i^* - \alpha_i\right) K(x_i, x) + b \qquad (25)$$

The author has respectively complied the polynomial kernel function, radial basis function, and Sigmoid kernel function Matlab procedures of SVM model.

Compiling Matlab grogram as follows:

Selecting the before 30 samples of sample set as training sample, the remaining 8 samples for the validation sample.

For output, +1 is the normal financial output, −1 is the financial crisis output.

Model parameters obtained by the methods of cross validation and lattice search.

Training the input model of training sample until obtains the satisfactory accuracy, this optimal classification function is the SVM model of construction industry financial risk.

3.3 Verifying the SVM Model of Construction Industry Financial Risk Evaluation

Respectively inputting training samples into above three kinds of SVM evaluation model of kernel function, testing the evaluation accuracy of training samples, then respectively inputting verification sample into model, it can verify the accuracy of the model for sample prediction, then evaluate the generalization ability of model. Obtaining the comprehensive accuracy of various model by calculating.

Comparing the model test results in Table 1 can be seen, the prediction accuracy of the three kernel model for training samples is the same, comprehensive prediction accuracy is generally high, so it is viable for SVM applying to construction industry financial risk evaluation. The prediction accuracy of radial basis kernel function model is higher than other two ones for testing samples. So Gauss radial basis kernel function model can better evaluate the construction industry financial risk.

Table 1 The model test results in

Samples	Accuracy		
	Sigmoid kernel function (%)	Polynomial kernel function (%)	Radial basis function (%)
Training samples	89.3	89.3	89.3
Verifying samples	80	80	90
Comprehensive	86.8	86.8	89.5

4 Conclusion

1. This paper has preprocessed the sample by cluster analysis and eliminated the inherent limitations of ST classification and Non ST, increased the seldom industry modeling possibilities of enterprises.
2. Trained samples using SVM theory, complied the polynomial kernel function, Sigmoid kernel function, radial basis function is superior to other kernel function ones, can better evaluate construction enterprise financial risk.
3. The choice of model parameters is difficult, such as penalty factor C, insensitive loss function ε, kernel parameter and so on. To optimize these parameters is worth further studying.

References

1. Zhou X, Li W (2008) The method and the model of financial early warning model at home and abroad. Finance Acc Mon (Integr Version) 5:79–84
2. Hui S, Wang W (2006) Financial early warning based on SVM study on the model and application. Comput Eng Des 27(7):1183–1186
3. Zhao X, Xue JL (2009) Forecasting financial status of listed company with data mining technique. China Manage Informationization 1(3):30–32
4. Xuegong Z (2000) Something about statistical learning theory and SVM. Acta Automatica Sin 26(1):32–42
5. Vapnik VN (1998) Statistical learning theory. Wiley, New York
6. Yao ZM (2006) The sensitivity analysis of artificially frozen soil temperature affecting factors. Hydrogeol Eng Geol 33(3):38–40
7. Liu J (2002) Training algorithm of SVM. Inf Control 31(1):45–50

Replaceable Encoding for Basic Variable of Traffic Network Transportation Optimization Problem

Kang Zhou, Yule Zhu, Feng Jin and Xin Tao

Abstract A linear programming problem is solved for the first time based on DNA computing model, which has important significance for research on DNA computing. According to feature of the mathematical model of traffic network transportation optimization problem (TNTOP), three groups of replaceable encoding for each basic variable are designed in the algorithm as follows: basic variable group; variable value group and c value group, which stores the information of basic variable and its value and has many groups replaceable foreign DNA corresponding to the basic variable. In the algorithm of TNTOP based on replaceable encoding for basic variable, combination operation which is used can assign or re-assign values to variables, which is composed of group insert experiment, gel electrophoresis experiment and group deleting experiment. The combination operation can test the constraint conditions and extract all optimal solutions of TNTOP. Detection experiment designed based on electrophoresis experiment can detect mixture containing many kinds of closed circle DNA sequence, and can detect out all closed circle DNA sequences or one closed circle DNA sequence of the mixture according to requirement of the algorithm. The correctness and the complexity of the DNA algorithm are proved, and a simulation example is given to explain feasibility of the DNA algorithm.

K. Zhou (✉) · Y. Zhu · F. Jin · X. Tao
School of Math and Computer, Wuhan Polytechnic University, 430023 Wuhan, China
e-mail: zhoukang_wh@yahoo.com.cn1

Y. Zhu
e-mail: 961220172@qq.com1

F. Jin
e-mail: 652836798@qq.com1

X. Tao
e-mail: 592100331@qq.com1

1 Introduction

With the urban traffic problem becoming increasingly serious, the problem becomes the bottleneck problem of the enterprises' development how to convey products safely and efficiently to utilize traffic network, study the problem has the important practical significance. In the research field of traffic network, optimization problems of traffic network have a variety of describing methods, and have all kinds of algorithms. However, DNA computing can not be applied to optimization problems of traffic network, the main reason is DNA computing can not solve linear programming problem at present.

In 2006, Doctor Zhou [1] put forward closed circle DNA computing model. Because of flexibility and diversity of closed circle DNA molecular structure, closed circle DNA computing model has solved plenty of practical problems [2–6], and application range of DNA computing is extended. How to further develop the performance of closed circle DNA encoding to solve more complex problems is an important research subject in closed circle DNA computing model. DNA algorithm of traffic network transportation optimization problem (TNTOP) is given by introducing replaceable encoding for basic variable into closed circle DNA computing model.

2 TNTOP

TNTOP is as follows: There are m warehouses in city A_i ($i = 1, 2, \ldots, m$) in traffic network, the quantity of goods which can be provided every day is a_i; there are n sales companys in city B_j ($j = 1, 2, \ldots, n$), the quantity of goods which will be sold every day is b_j, unit transportation price from city A_i to city B_j is c_{ij}. Traffic and transportation scheme every day is determined to meet the need for the quantity of goods, which makes total transportation price reach the minimum value.

For simplicity, let the total quantity of goods provided be exactly equal to the total quantity of goods sold; and let a_i, b_j and c_{ij} be positive integers. Linear programming model of TNTOP is as follows:

$$\min z = \sum \sum \left(c_{ij} \times x_{ij} | 1 \leq i \leq m; \ 1 \leq j \leq n \right).$$

s.t.

$$\sum (x_{ij} | 1 \leq i \leq m) = b_j, \quad j = 1, 2, \ldots, n. \tag{1}$$

$$\sum (x_{ij} | 1 \leq j \leq n) = a_i, \quad i = 1, 2, \ldots, m. \tag{2}$$

$$x_{ij} \geq 0.$$

3 Replaceable Encoding for Basic Variable

To improve closed circle DNA computing model [1] in order to solve practical problems is the important dynamics of development of DNA computing. For $n \times m$ decision variables, $n \times m$ closed circle DNA is too long led to decrease of the success rate of biochemistry experiment. $n + m - 1$ closed circle DNA can represent $n + m - 1$ basic variables of this model. The new characteristics of closed circle DNA encoding are as follows:

1. The $n + m + 1$ closed circle DNA is formed to arrange DNA encoding of the base variable according to a certain order, so DNA encoding in the same position for different closed circle DNA molecules can be different.
2. Each DNA encoding must recognize city A_i, city B_j and traffic, so encoding of closed circle DNA can recognize all base variables and their values.
3. Closed circle DNA has three functions as follows: it can do solution feasible operation; it can recognize all optimal solutions; it can be detected all DNA encoding. They request closed circle DNA encoding to be replaceable.

According to characteristics of the DNA encoding, recognition sequence in the h bit of closed circle DNA is r_{ijk}, this means $x_{ij} = k$; for recognition sequence r_{ijk}, two groups of encodings of foreign DNA are designed as p_{ijk} and q_{ijk}, whose length means respectively that $x_{ij} = k$ and that transportation price of x_{ij} is $c_{ij} \times k$. For closed circle DNA encoding $r_{i(1)j(1)k(1)} \quad r_{i(2)j(2)k(2)} \ldots r_{i(n+m+1)j(n+m+1)k(n+m+1)}$, arrangement rule is that $i(h) < i(h + 1)$ or when $i(h) = i(h + 1)$, $j(h) < j(h + 1)$. Only one group of foreign DNA encodings can insert into recognition sequence r_{ijk}, so two groups of foreign DNA encodings of closed circle DNA can switch into another one. The two groups of foreign DNA encodings are called replaceable encoding for basic variable.

Extended closed circle DNA computing model has group insert experiment and group delete experiment.

Group insert experiment: Foreign DNA molecular $p_{ijk}(k = 1, 2, \ldots, K)$ is inserted into closed circle DNA molecular.
Group delete experiment: Foreign DNA molecular $p_{ijk}(k = 1, 2, \ldots, K)$ is removed from all closed circle DNA molecular in mixture.

4 DNA Algorithm of TNTOP

4.1 Closed Circle DNA Algorithm

Closed circle DNA algorithm of TNTOP is as follows:

Step 1 Encode replaceable encodings for basic variable corresponding to basic variable $x_{ij}(i = 1, 2, \ldots, m; j = 1, 2, \ldots, n)$ and $k = 1, 2,\ldots, \min[a_i,b_j]$ is as follows:

The first (basic variable) group: $3'-r_{ij0}-5'$, $3'-r_{ijk}-5' = s_{ijk}t_{ijk}$, $3'-r_0-5'$, where $|r_{ij0}| = |r_{ijk}| = 15$ bp, $|r_0| = 6$ bp;

The second (variable value) group: $3'-p_{ijk}-5' = t_{ijk}u_{ijk}s_{ijk}$, where $|p_{ijk}| = 15 \times k$ bp;

The third (c value) group: $3'-q_{ijk}-5' = t_{ijk}v_{ijk}s_{ijk}$, where$|q_{ijk}| = 15 \times c_{ij} \times k$ bp;

Where the encoding r_{ij0} of the first group represents $x_{ij} = 0$, and the encoding r_{ijk} represents $x_{ij} = k$, r_{ij0} are not recognition sequences, $s_{ijk}t_{ijk}$ and r_0 are recognition sequences; the encoding of the second group represents value of x_{ij} being k; the encoding of the third group represents transport price of x_{ij} being $c_{ij} \times x_{ij}$.

Step 2 Construct all initial solution of TNTOP.

1. For $i, i(1), i(2) = 1, 2, \ldots, m; j, j(1), j(2) = 1, 2, \ldots, n; k = 0, 1, \ldots, \min[a_i, b_j]$; $k(1) = 0, 1, \ldots, \min[a_{i(1)}, b_{j(1)}]; k(2) = 0, 1, \ldots, \min[a_{i(2)}, b_{j(2)}]$, to be notice **that $i(1) < i(2)$ or when $i(1) = i(2), j(1) < j(2)$.**

First synthesize DNA fragments of the first group, then synthesize four groups of DNA fragments as follows:

A group $3'-r_0r_{ijk}-5'$;

B group $3'-r_{i(1)j(1)k(1)}r_{i(2)j(2)k(2)}-5'$;

C group $5'-r'_{i(1)j(1)k(1)}r'_{i(2)j(2)k(2)}-3'$;

D group $5'-r'_{ijk}r'_0-3'$

2. Construct data pool by hybridization reaction as follows:
(a) Renaturation reaction for A group; let $k = 1$.
(b) Hybridization reaction for B group and new mixture, and DNA fragments of no more than $15(n + m + 1) + 6$ bp chain length are saved by electrophoresis experiment.
(c) If $k \geq n + m - 1$, then turn to e; Otherwise, let $k = k + 1$, turn to d.
(d) Hybridization reaction for C group and new mixture, and DNA fragments of no more than $15(n + m + 1) + 6$ bp chain length are saved by electrophoresis experiment. let $k = k + 1$, turn to b.
(e) Hybridization reaction for D group and new mixture, and DNA fragments of no more than $15(n + m + 1) + 6$ bp chain length are saved by electrophoresis experiment.
(f) Cyclization reaction for new mixture, and DNA fragments of $L = 15(n + m + 1) + 6$ bp chain length are saved by electrophoresis experiment.

Step 3 Filter out all basic feasible solutions of TNTOP.

1. Let $j = 1$.
2. Let $i = 1$.
3. Group insert experiment for foreign DNA molecular p_{ijk}.

4. If $i = m$, then turn to (5); Otherwise, let $i = i + 1$, turn to (3).
5. Closed circle DNA of $L_j = 15(n + m + b_j + 1) + 6$ bp chain length are saved by electrophoresis experiment.
6. Let $i = 1$.
7. Group delete experiment for foreign DNA molecular p_{ijk}.
8. If $i = m$, then turn to (9); Otherwise, let $i = i + 1$, turn to (7).
9. If $j = n$, then turn to (10); Otherwise, let $j = j + 1$, turn to (2).
10. Let $i = 1$.
11. Let $j = 1$.
12. Group insert experiment for foreign DNA molecular p_{ijk}.
13. If $j = n$, then turn to (14); Otherwise, let $j = j + 1$, turn to (12).
14. Closed circle DNA of $L'_i = 15(n + m + a_i + 1) + 6$ bp chain length are saved by electrophoresis experiment.
15. Let $j = 1$.
16. Group delete experiment for foreign DNA molecular p_{ijk}.
17. If $j = n$, then turn to (18); Otherwise, let $j = j + 1$, turn to (16).
18. If $i = m$, then turn to Step 4; Otherwise, let $i = i + 1$, turn to (11).

Step 4 Filter out all optimal solutions of TNTOP.

1. Let $j = 1$.
2. Let $i = 1$.
3. Group insert experiment for foreign DNA molecular q_{ijk}.
4. If $i = m$, then turn to (5); Otherwise, let $i = i + 1$, turn to (3).
5. If $j = n$, then turn to (6); Otherwise, let $j = j + 1$, turn to (2).
6. All the shortest DNA fragments are saved by electrophoresis experiment.
7. Let $j = 1$.
8. Let $i = 1$.
9. Group delete experiment for foreign DNA molecular q_{ijk}.
10. If $i = m$, then turn to (11); Otherwise, let $i = i + 1$, turn to (9).
11. If $j = n$, then turn to Step 5; Otherwise, let $j = j + 1$, turn to (8).

Step 5 Detect all optimal solutions of TNTOP.

1. Chain length of closed circle DNA $L_{10} = 15(n + m + 1) + 6$ bp, let $i = 1$.
2. Let $j = 1$.
3. Group insert experiment for foreign DNA molecular p_{ijk}.
4. All the shortest DNA fragments are saved by electrophoresis experiment. Chain length of closed circle DNA $L_{ij} = L_{i(j-1)} + 15 \times k$ bp, then $x_{ij} = k$.
5. If $j = n$, then turn to (6); Otherwise, let $j = j + 1$, turn to (3).
6. If $i \leq m - 1$, then let $L_{in} = L_{(i+1)0}$, $i = i + 1$, turn to (2); Otherwise, detect all optimal solutions of TNTOP.

4.2 Analysis on Correctness and Complexity of DNA Algorithm

Theorem 1 (Theorem of correctness of algorithm) *Any one closed circle DNA of Step 4 represents optimal solution of TNTOP. Closed circle DNA mixture of Step 4 contains closed circle DNA which represents optimal solution of TNTOP.*

Proof Closed circle DNA mixture of Step 2 has the properties as follows:

1. Any one closed circle DNA can be expressed as $(r_{i(1)j(1)k(1)}r_{i(2)j(2)k(2)}\cdots r_{i(n+m+1)}$ $_{j(n+m+1)k(n+m+1)})$;
2. According to B group and C group, arrangement rule is that $i(h) < i(h+1)$ or when $i(h) = i(h+1)$, $j(h) < j(h+1)$, for $\forall h = 1, 2,\ldots, n+m$.
3. According to A group and parallelism of hybridization reaction, The first group of all permutation of $n+m+1$ can be obtained by Step 2.

So closed circle DNA mixture of Step 2 represents set of all initial solution which contains all basic feasible solutions of TNTOP.

For $j = 1$ of Step 3, chain length of closed circle DNA after (4) is $15[n+m+\sum(k(h)|j(h) = 1\wedge 1 \le h \le n+m+1)+1]+6$ bp. So closed circle DNA of (5) whose chain length is $15(n+m+b_1+1)+6$ bp satisfies constraint (1) of $j = 1$. Closed circle DNA after (9) satisfies constraint (1). And closed circle DNA after (18) satisfies constraint (1) and (2). So closed circle DNA mixture of Step 3 represents set of all basic feasible solutions of TNTOP.

For Step 4, chain length of closed circle DNA after (5) is $15\left[n+m+\sum\left(c_{i(h)j(h)} \times k(h)|1 \le h \le n+m+1\right)+1\right]+6$ bp. the shortest DNA whose chain length is $15\sum\left(c_{i(h)j(h)} \times k(h)|1 \le h \le n+m+1\right)$, so closed circle DNA mixture of Step 4 represents set of all optimal solutions of TNTOP.   □

Theorem 2 *For m warehouses which can be provided every day is a_i and n sales companys which will be sold every day is b_j, unit transportation price from city A_i to city B_j is c_{ij}. Let $L = max\ min[a_i,b_j]$, then DNA algorithm of TNTOP needs at most $5 \times m \times n \times L$ DNA encodings, 5 test tubes, and operating times of DNA algorithm is $8nm + 5n + 3m + 13$.*

Proof For Step 1, the first group needs at most $m \times n \times L$ DNA encodings; the second group and the third group need respectively at most $2 \times m \times n \times L$ DNA encodings, so DNA algorithm of TNTOP needs at most $5 \times m \times n \times L$ DNA encodings.

Step 2 needs 5 test tubes, and does $3 + 3 \times (n+m+1) + 6$ experiments; Step 3 needs 2 test tubes, and does $2 \times n \times (m+1+m)$ experiments; Step 4 needs 2 test tubes, and does $n \times m + 1 + n \times m$ experiments; Step 5 needs 2 test tubes, and does $2 \times n \times m$ experiments, so DNA algorithm of TNTOP needs 5 test tubes, and operating times of DNA algorithm is $8nm + 5n + 3m + 13$.   □

5 Analysis on Validity of DNA Algorithm

We do a simulation experiment to verify validity and correctness of closed circle DNA algorithm of TNTOP. TNTOP is as follows: There are 3 warehouses in city A_i ($i = 1, 2, 3$) in traffic network$_i$; there are 2 sales companys in city B_j ($j = 1, 2$), the provided quantity, the sold quantity and unit transportation price from city A_i to city B_j see Table 1.

Step 1 Compute the value range of decision variables as follows:

$$0 \le x_{11} \le 1; \quad 0 \le x_{12} \le 2; \quad 0 \le x_{21} \le 1;$$
$$0 \le x_{22} \le 1; \quad 0 \le x_{31} \le 1; \quad 0 \le x_{32} \le 1.$$

The first (basic variable) group:

r_{110} AAACTATTCATTCCC;
r_{111} CTATCTATC-AACGTT, corresponding to restriction endonuclease: Acl I;
r_{120} CCTCACATTAATATC;
r_{121} CTTATCCTC-CTGAAG, corresponding to restriction endonuclease: Acu I;
r_{122} CTTATCATC-ACCGGT, corresponding to restriction endonuclease: Age I;
r_{210} AACTAATTCATCCTC;
r_{211} CCCATTTTC-CTTAAG, corresponding to restriction endonuclease: Afl II;
r_{220} AAACTTACTATTCCC;
r_{221} TACATCCTC-ATTAAT, corresponding to restriction endonuclease: Ase I;
r_{310} TTTACAAACACTCTC;
r_{311} TCAATCCCT-CCTAGG, corresponding to restriction endonuclease: Avr II;
r_{320} AAACTTCATACTTCC;
r_{321} CTAACTCTC-GGATCC, corresponding to restriction endonuclease: BamH I;
r_0 GTGCAC, corresponding to restriction endonuclease: ApaL I.

The second (variable value) group:

u_{111} ATACTTATA;
u_{121} TACTCTATT;
u_{122} AACTAATCTATCCCTAACTAACTTATTCCC;
u_{211} CATTTATTC;
u_{221} TCTTATTCA;
u_{311} CTTTCACCA;

Table 1 Data of TNTOP

Unit transportation price		Sales company			The sold quantity
		v_1	v_2	v_3	
Warehouse	u_1	2	1	2	1
	u_2	1	2	1	3
The provided quantity		2	1	1	

u_{321} ACTAAATTC.

The third (c value) group:
v_{111} CTTCCTAACTTCTCTAATACACAC;
v_{121} CTCCTTATA;
v_{122} TATATCATCCCTAATCATACTCCT;
v_{211} CTCCACATT;
v_{221} TACATTACTCATTTTATCACACAC;
v_{311} TATCATATCTTAAACCTTATTACC;
v_{321} CATTTAACC.

Step 2 Construct 96 initial solution of TNTOP as follows:

$(r_0 r_{110} r_{120} r_{210} r_{220} r_{310} r_{320})$; $(r_0 r_{111} r_{120} r_{210} r_{220} r_{310} r_{320})$; $(r_0 r_{110} r_{121} r_{210} r_{220} r_{310} r_{320})$;
$(r_0 r_{111} r_{121} r_{210} r_{220} r_{310} r_{320})$;...;$(r_0 r_{111} r_{122} r_{211} r_{221} r_{311} r_{321})$.

Step 3 Filter out all basic feasible solutions of TNTOP as follows:

1. When $j = 1$, filter out 36 closed circle DNA as follows:

$(r_0 r_{111} r_{120} r_{210} r_{220} r_{310} r_{320})$;
$(r_0 r_{110} r_{120} r_{211} r_{220} r_{310} r_{320})$;...;$(r_0 r_{110} r_{122} r_{210} r_{221} r_{311} r_{321})$.

2. When $j = 2$, filter out 9 closed circle DNA as follows:

$(r_0 r_{111} r_{122} r_{210} r_{221} r_{310} r_{320})$; $(r_0 r_{111} r_{122} r_{210} r_{220} r_{310} r_{321})$; $(r_0 r_{111} r_{121} r_{210} r_{221} r_{310} r_{321})$;
$(r_0 r_{110} r_{122} r_{211} r_{221} r_{310} r_{320})$; $(r_0 r_{110} r_{122} r_{211} r_{220} r_{310} r_{321})$; $(r_0 r_{110} r_{121} r_{211} r_{221} r_{310} r_{321})$;
$(r_0 r_{110} r_{122} r_{210} r_{221} r_{311} r_{320})$; $(r_0 r_{110} r_{122} r_{210} r_{220} r_{311} r_{321})$; $(r_0 r_{110} r_{121} r_{210} r_{221} r_{311} r_{321})$.

3. When $i = 1$, filter out 5 closed circle DNA as follows:

$(r_0 r_{111} r_{121} r_{210} r_{221} r_{310} r_{321})$; $(r_0 r_{110} r_{122} r_{211} r_{221} r_{310} r_{320})$; $(r_0 r_{110} r_{122} r_{211} r_{220} r_{310} r_{321})$;
$(r_0 r_{110} r_{122} r_{210} r_{221} r_{311} r_{320})$; $(r_0 r_{110} r_{122} r_{210} r_{220} r_{311} r_{321})$.

4. When $i = 2$, filter out 3 closed circle DNA as follows:

$(r_0 r_{111} r_{121} r_{210} r_{221} r_{310} r_{321})$; $(r_0 r_{110} r_{122} r_{211} r_{220} r_{310} r_{321})$;
$(r_0 r_{110} r_{122} r_{210} r_{221} r_{311} r_{320})$.

5. Basic feasible solutions are as follows:

$(r_0 r_{111} r_{121} r_{210} r_{221} r_{310} r_{321})$; $(r_0 r_{110} r_{122} r_{211} r_{220} r_{310} r_{321})$;
$(r_0 r_{110} r_{122} r_{210} r_{221} r_{311} r_{320})$.

Step 4 Filter out all optimal solutions of TNTOP as follows:

$$(r_0 r_{110} r_{122} v_{122} r_{122} r_{211} v_{211} r_{211} r_{220} r_{310} r_{321} v_{321} r_{321}), \text{Chain length is 156bp.}$$

Step 5 Detect all optimal solutions of TNTOP.

Optimal solution of TNTOP is $X^* = (0, 2, 1, 0, 0, 1)T$, and optimal solution value is $Z^* = (|r_0 r_{110} r_{122} v_{122} r_{122} r_{211} v_{211} r_{211} r_{220} r_{310} r_{321} v_{321} r_{321}| - 96)/15$.

References

1. Kang Z, Jin X (2006) Closed circle DNA algorithm of the minimal covering problem. Comput Eng Appl 20:7–9
2. Kang Z, Shuo L, Lei Q et a1. (2011) The study of plasmid DNA computing model. J HuaZhong Univ Sci Technol (Nat Sci) 2:32–35
3. Kang Z, Zhi X, Chanjia W et a1. (2010) Closed circle DNA algorithm of set covering problem. J HuaZhong Univ Sci Technol (Nat Sci) 2:21–25
4. Kang Z, Lei Q, Xiaojun T et al (2009) Closed circle DNA algorithm of 0–1 planning problem. Syst Eng Electron 4:947–951
5. Kang Z, Xiaojun T, Jin X (2009) Closed circle DNA algorithm of change positive- weighted Hamilton circuit problem. J Syst Eng Electron 3:636–642
6. Kang Z, Chuanjia W, Shuo L et a1. (2009) Closed circle DNA algorithm for SAT problem. Huazhong Keji Daxue Xuebao (Ziran Kexue Ban) 7:75–78

Upper and Lower Solutions of Boundary Value Problems for Impulsive Fractional Differential

Chuan xia Hou and Yong Li

Abstract This paper studies the existence of solutions for boundary value problem of impulsive fractional differential equations, we obtain the method of upper and lower solutions by using Schauder's fixed point theorem.

Keywords Fractional differential equation · Boundary value problem · Impluse · Upper and lower solution · Fixed point theorem

1 Introduction

In this paper, we consider the following implusive fractional differential equation

$$\begin{cases} {}^{c}D_{0^{+}}^{q}x(t) + f(t,x) = 0, & 1 < q \le 2, \quad 0 < t < 1, \quad t \ne t_1, \\ \Delta x|_{t=t_1} = I(x(t_1)), & \Delta x'|_{t=t_1} = -\frac{1}{1-t_1}I(x(t_1)), \\ x(0) = x(1) = 0. \end{cases} \quad (1.1)$$

where ${}^{c}D_{0^{+}}^{q}$ is the Caputo fractional derivative, $\boldsymbol{f(t,x)} \in C((0,1) \times (0,+\infty)$, $[0,+\infty))$, $f(t,1) \ne 0, t \in (0,1)$. $\Delta x|_{t=t_1} = x(t_1^+) - x(t_1), \Delta x'|_{t=t_1} = x'(t_1^+) - x'(t_1)$, and I is nondecreasing in x.

Recently, much attention has been focused on the study of solutions of boundary value problem for fractional differential equations, see [1–6]. However,

C. x. Hou (✉)
School of Mathematical sciences, University of Jinan, Jinan 250022 Shandong, People's Republic of China
e-mail: ss_houcx@ujn.edu.cn

Y. Li
University of Jinan, No. 6 middle School of Licheng, Jinan 250100 Shandong, People's Republic of China
e-mail: 21047985@qq.com

impulsive differential equations of fractional order have not been much studied, for some recent work on impulsive fractional differential equations, see [7–12] and the references therein.

2 Preliminaries and Lemmas

Let us recall some basic definitions and some known results.

Definition 1 [4]   The Caputo fractional derivative of order q>0 of a continuous function $f : (0, +\infty) \to R$ is given by

$$^c D_{0^+}^q f(t) = \frac{1}{\Gamma(n-q)} \int_0^t \frac{f^{(n)}(s)}{(t-s)^{q-n+1}} ds,$$

where n is the smallest integer greater than or equal to q, provided that the right side is pointwise defined on $(\boldsymbol{0}, +\infty)$.

Definition 2 [4]   The Riemann–Liouville fractional integral of order $q > 0$ of a function $f : (0, +\infty) \to R$ is given by

$$I_{0^+}^q f(t) = \frac{1}{\Gamma(q)} \int_0^t (t-s)^{q-1} f(s) ds,$$

provided that the right side is pointwise defined on $(\boldsymbol{0}, +\infty)$.

Lemma 1 [4]   *Let $q > 0$. Assume $u \in C^n[0, 1]$. Then*

$$I_{0^+}^q {}^c D_{0^+}^q u(t) = u(t) + C_0 + C_1 t + \cdots + C_{n-1} t^{n-1},$$

for some $C_i \in R, i = 1, 2, \cdots, n-1$, where n is the smallest integer greater than or equal to q.

   Let $PC[[0, 1], R^+] = \{x | x : J \to R, x(t)$ is   continuous   except   at $t = t_1, x(t_1^+), x(t_1^-)$ exist and $x(t_1) = x(t_1^-)\}$. Let $J' = [0, 1] \backslash \{t_1\}$. Now we introduce the following definition about the upper and lower solutions of problem (1.1).

Definition 3 Let $\alpha \in PC[[0, 1], R^+] \cap C^2[J', R^+]$ satisfying

$$\begin{cases} {}^c D_{0^+}^q \alpha(t) + f(t, \alpha(t)) \geq 0, & 1 < q \leq 2, \quad 0 < t < 1, \quad t \neq t_1, \\ \Delta \alpha|_{t=t_1} = I(\alpha(t_1)), \quad \Delta \alpha'|_{t=t_1} = -\frac{1}{1-t_1} I(\alpha(t_1)), \\ \alpha(0) \leq 0, \quad \alpha(1) \leq 0, \end{cases}$$

Then $\alpha(t)$ is called a lower solution of (1.1). Similarly, we can define an upper solution of (1.1) visa changing the symbol of that inequality.

Lemma 2 *For a given $y \in C[0, 1]$, a function x is a solution of the impulsive boundary value problem*

$$\begin{cases} {}^cD_{0+}^q x(t) + y(t) = 0, & 1 < q \leq 2, \quad 0 < t < 1, \quad t \neq t_1, \\ \Delta x|_{t=t_1} = I(x(t_1)), \quad \Delta x'|_{t=t_1} = -\frac{1}{1-t_1} I(x(t_1)), \\ x(0) = x(1) = 0, \end{cases} \tag{2.1}$$

if and only if x is a solution of the impulsive fractional integral equation

$$x(t) = \begin{cases} -\frac{1}{\Gamma(q)} \int_0^t (t-s)^{q-1} y(s)ds - c_2 t, & 0 < t \leq t_1, \\ -\frac{1}{\Gamma(q)} \int_{t_1}^t (t-s)^{q-1} y(s)ds - \frac{1}{\Gamma(q)} \int_0^{t_1} (t_1-s)^{q-1} y(s)ds \\ \quad -\frac{t-t_1}{\Gamma(q-1)} \int_0^{t_1} (t_1-s)^{q-2} y(s)ds + (1 - \frac{t-t_1}{1-t_1})I(x(t_1)) - c_2 t, & t_1 < t \leq 1, \end{cases}$$
$$\tag{2.2}$$

where

$$c_2 = -\frac{1}{\Gamma(q)} \int_{t_1}^1 (1-s)^{q-1} y(s)ds - \frac{1}{\Gamma(q)} \int_0^{t_1} (t_1-s)^{q-1} y(s)ds$$
$$- \frac{1-t_1}{\Gamma(q-1)} \int_0^{t_1} (t_1-s)^{q-2} y(s)ds.$$

Proof Suppose x is a solution of (2.1). Using Lemma 1, for some constants $c_1, c_2 \in R$, we have

$$x(t) = -\frac{1}{\Gamma(q)} \int_0^t (t-s)^{q-1} y(s)ds - c_1 - c_2 t, \quad 0 < t \leq t_1. \tag{2.3}$$

Then we have

$$x'(t) = -\frac{1}{\Gamma(q-1)} \int_0^t (t-s)^{q-2} y(s)ds - c_2, \quad 0 < t \leq t_1. \tag{2.4}$$

If $t_1 < t \leq 1$, for some constants $d_1, d_2 \in R$, we have

$$x(t) = -\frac{1}{\Gamma(q)} \int_{t_1}^t (t-s)^{q-1} y(s)ds - d_1 - d_2(t - t_1),$$

$$x'(t) = -\frac{1}{\Gamma(q-1)} \int_{t_1}^t (t-s)^{q-2} y(s)ds - d_2,$$

In view of $\Delta x|_{t=t_1} = I(x(t_1))$, $\Delta x'|_{t=t_1} = -\frac{1}{1-t_1} I(x(t_1))$, we have

$$-d_1 = -\frac{1}{\Gamma(q)} \int_0^{t_1} (t_1-s)^{q-1} y(s)ds + I(x(t_1)) - c_1 - c_2 t_1,$$

$$-d_2 = -\frac{1}{\Gamma(q-1)} \int_0^{t_1} (t_1-s)^{q-2} y(s)ds - \frac{1}{1-t_1} I(x(t_1)) - c_2.$$

Thus, we obtain

$$x(t) = -\frac{1}{\Gamma(q)}\int_{t_1}^{t}(t-s)^{q-1}y(s)ds - \frac{1}{\Gamma(q)}\int_{0}^{t_1}(t_1-s)^{q-1}y(s)ds + (1-\frac{t-t_1}{1-t_1})I(x(t_1))$$
$$-\frac{t-t_1}{\Gamma(q-1)}\int_{0}^{t_1}(t_1-s)^{q-2}y(s)ds - c_1 - c_2 t, \quad t_1 < t < 1. \tag{2.5}$$

By the boundary value condition $x(0) = 0$ and (2.3),(2.4),we have $c_1 = 0$. By condition $x(1) = 0$ and (2.5), we obtain

$$c_2 = -\frac{1}{\Gamma(q)}\int_{t_1}^{1}(1-s)^{q-1}y(s)ds - \frac{1}{\Gamma(q)}\int_{0}^{t_1}(t_1-s)^{q-1}y(s)ds$$
$$-\frac{1-t_1}{\Gamma(q-1)}\int_{0}^{t_1}(t_1-s)^{q-2}y(s)ds.$$

Substituting the value c_1, c_2 in (2.3) and (2.5), we obtain (2.2). Conversely, we assume that x is a solution of (2.2), by a direct computation, it follows that the solution given by (2.2) satisfies (2.1).

3 Main Results

Theorem 1 *Let α and β respectively be a lower and upper solution of* (1.1) *on [0,1] satisfying*

$$(a_1) \quad \alpha(t) \le \beta(t), t \in [0, 1],$$
$$(a_2) \quad U_\alpha^\beta = \{(t,x) \in (0,1) \times R : \alpha(t) \le x(t) \le \beta(t)\},$$
$$(a_3) \quad |f(t,x)| \le h(t), (t,x) \in U_\alpha^\beta, \quad h \in C((0,1), R^+).$$

If $\frac{1}{\Gamma(q-1)}\int_0^1(1-s)^{q-2}h(s)ds < \infty$, then problem (1.1) *has at least a solution $x(t) \in PC([0,1], R^+)$, and $\alpha(t) \le x(t) \le \beta(t)$.*

Proof Let

$$f^*(t,x) = \begin{pmatrix} f(t,\alpha(t)) + (\alpha(t)-x)^{\frac{1}{2}}, & x < \alpha(t), \\ f(t,x(t)), & \alpha(t) \le x \le \beta(t), \\ f(t,\beta(t)) - (x-\beta(t))^{\frac{1}{2}}, & x > \beta(t). \end{pmatrix}$$

By (a_2) and the definition of f^*, we know: $f^* : (0,1) \times (0,\infty) \to [0,\infty)$, and from (a_3) we know $|f^*(t,x)| \le h(t)$, $(t,x) \in (0,1) \times R$. Let

$$I^*(x) = \begin{pmatrix} I(\alpha(t)), & x < \alpha(t), \\ I(x(t)), & \alpha(t) \le x \le \beta(t), \\ I(\beta(t)), & x > \beta(t). \end{pmatrix}$$

Then $I^* \in C(R, R)$. Consider the following boundary value problems

$$
\begin{cases}
{}^cD_{0+}^q x(t) + f(t, x) = 0, & 1 < q \le 2, \quad 0 < t < 1, \quad t \ne t_1, \\
\Delta x|_{t=t_1} = I(x(t_1)), \quad \Delta x'|_{t=t_1} = -\frac{1}{1-t_1} I(x(t_1)), \\
x(0) = x(1) = 0
\end{cases}
\tag{3.1}
$$

The following we will show that if $x(t) \in PC([0, 1], R^+)$ is the solution of (3.1), then $\alpha(t) \le x \le \beta(t)$. Let $z(t) = \alpha(t) - x(t)$. We will to prove $z(t) \le 0$. If not, then exist some points $t' \in [0, 1]$ satifies $z(t) > 0$. There have two possibilities:

case 1: $z(t_1) > 0$. Then $I^*(x(t_1)) = I(\alpha(t_1))$, and $z(t_1^+) = z(t_1)$, so $z(t)$ is continuous on [0,1], thus $z(t)$ has a maximum point $s \in [0, 1]$, and $s \ne 0, s \ne 1$, then s have three cases:

i. $s < t_1$, that is $s \in (0, t_1)$, then $z'(s) = 0$ and ${}^cD_{0+}^q z(s) \le 0$. But

$$
{}^cD_{0+}^q x(s) = -f^*(s, x(s)) = -\left[f(s, \alpha(s)) + (\alpha(s) - x)^{\frac{1}{2}}\right] < -f(s, \alpha(s)) \le {}^cD_{0+}^q \alpha(s),
$$

this contradict with ${}^cD_{0+}^q z(s) \le 0$.

ii. $s > t_1$, similar to the discussion of (i) can prove.

iii. $s = t_1$. Then $f^*(s, x(s)) = f(s, \alpha(s))$, $\psi I^*(x(s)) = I(a(s))$, so s is a continuous point of $z(s)$, then $z'(s) = 0$ and ${}^cD_{0+}^q z(s) \le 0$. But

$$
{}^cD_{0+}^q x(t_1^-) = \lim_{t \to t_1^-} {}^cD_{0+}^q x(t) = -\lim_{t \to t_1^-} f^*(t, x(t))
$$
$$
= -\left[f(t_1, \alpha(t_1)) + (\alpha(t_1) - x)^{\frac{1}{2}}\right] < -f(t_1, \alpha(t_1)) \le {}^cD_{0+}^q \alpha(t_1^-).
$$

Thus ${}^cD_{0+}^q z(t_1^-) > 0$. Similarly we have ${}^cD_{0+}^q z(t_1^+) > 0$. This contradict with ${}^cD_{0+}^q z(s) \le 0$.

case 2: $z(t_1) \le 0$. We first prove $z(t) \le 0$ on $[0, t_1]$. If not, then $z(t)$ has a maximum value on $[0, t_1]$, by (i) we obtain contradiction. In veiw of $z(t_1) \le 0$ and the increasing of I, we have

$$
\alpha(t_1) + I(\alpha(t_1)) \le x(t_1) + I(x(t_1)),
$$

this namely $z(t_1^+) \le 0$. The following we prove $z(t) \le 0$ on $(t_1, 1]$ If not, then there exist $s \in (t_1, 1]$ such that $z(s) > 0$ and exist $t^* \in (t_1, s)$ such that $z(s) > z(t^*) = 0$, then $z(t)$ has a maximum value on $[t^*, 1]$, by (ii) we obtain contradiction. To sum up we have $x(t) \ge \alpha(t)$. Similar to prove $x(t) \le \beta(t)$. The following we will prove there at least a solution of (3.1). Set

$$T(x)(t) = \begin{cases} -\frac{1}{\Gamma(q)}\int_0^t (t-s)^{q-1}f^*(s,x(s))ds - Mt, & 0<t\le t_1, \\ -\frac{1}{\Gamma(q)}\int_{t_1}^t (t-s)^{q-1}f^*(s,x(s))ds - \frac{1}{\Gamma(q)}\int_0^{t_1}(t_1-s)^{q-1}f^*(s,x(s))ds \\ \quad -\frac{t-t_1}{\Gamma(q-1)}\int_0^{t_1}(t_1-s)^{q-2}f^*(s,x(s))ds + (1-\frac{t-t_1}{1-t_1})I(x(t_1)) - Mt, & t_1<t\le 1, \end{cases}$$

where

$$M = -\frac{1}{\Gamma(q)}\int_{t_1}^1 (1-s)^{q-1}f^*(s,x(s))ds - \frac{1}{\Gamma(q)}\int_0^{t_1}(t_1-s)^{q-1}f^*(s,x(s))ds$$
$$-\frac{1-t_1}{\Gamma(q-1)}\int_0^{t_1}(t_1-s)^{q-2}f^*(s,x(s))ds.$$

Then $T : X \to X = PC([0,1]R)$. Clearly, the fixed points of operator T are the solutions of problem (3.1). By the condition

$$\frac{1}{\Gamma(q-1)}\int_0^1 (1-s)^{q-2}h(s)ds < \infty,$$

we know TX is bounded. For $0<t\le t_1$,

$$\left|\frac{d}{dt}T(x)(t)\right| = \left|-\frac{1}{\Gamma(q-1)}\int_0^t (t-s)^{q-2}f^*(s,x(s))ds - M\right|$$
$$\le \frac{2}{\Gamma(q-1)}\int_0^1 (1-s)^{q-2}|f^*(s,x(s))|ds + \frac{2}{\Gamma(q)}\int_0^1 (1-s)^{q-1}|f^*(s,x(s))|ds$$
$$\le \frac{4}{\Gamma(q-1)}\int_0^1 (1-s)^{q-2}h(s)ds < \infty.$$

If $t_1<t<1$,

$$\left|\frac{d}{dt}T(x)(t)\right| = \left|-\frac{1}{\Gamma(q-1)}\int_{t_1}^t (t-s)^{q-2}f^*(s,x(s))ds\right.$$
$$\left. -\frac{1}{\Gamma(q-1)}\int_0^{t_1}(t_1-s)^{q-2}f^*(s,x(s))ds - \frac{1}{1-t_1}I^*(x(t_1)) - M\right|$$
$$\le \frac{5}{\Gamma(q-1)}\int_0^1 (1-s)^{q-2}h(s)ds + \frac{1}{1-t_1}I^*(x(t_1)) < \infty,$$

which implies that TX is relative compact, as a consequence of Schauder's fixed point theorem, we deduce that T has a fixed point respectively in $0<t\le t_1$ and $t_1 < t < 1$ which is the solution of problem (3.1). So there at least exit a solution $x(t) \in PC([0,1],R^+)$ of (3.1), and $\alpha(t) \le x(t) \le \beta(t)$.

References

1. Zhang S (2006) Positive solutions for boundary-value problems of nonlinear fractional differential equations. Electron J Differ Equ 36:1–12
2. Qiu T, Bai Z (2008) Existence of positive solutions for singular fractional equations. Electron J Differ Equ 146:1–9
3. Bai Z, Lu H (2005) Positive solutions for boundary value problem of nonlinear fractional equation. J Math Anal Appl 311:495–505
4. Kilbas AA, Srivastava HH, Trujillo JJ (2006) Theory and applications of fractional differential equations. Elsevier Science B. V, Amsterdam
5. Wang G, Ntouyas SK, Zhang L (2011) Positive solutions of the three-point boundary value problem for fractional-order differential equations with an advanced argument. Adv Differ Equ 2:1–16
6. Yang X, Wei Z, Dong W (2012) Existence of positive solutions for the boundary value problem of nonlinear fractional differential equations. Commun Nonlinear Sci Numer Simulat 17:85–92
7. Ahmad B, Sivasundaram S (2009) Existence results for nonlinear impulsive hybird boundary value problems involving fractional differential equations. Nonlinear Anal 3:251–258
8. Benchohra M, Hamani S (2009) The method of upper and lower solutions and impulsive fractional differential inclusions. Nonlinear Anal 3:433–440
9. Tian Y, Bai Z (2010) Existence results for the three-point impulsive boundary value problem involving fractional differential equations. Comput Math Appl 59:2601–2609
10. Wang G, Ahmad B, Zhang L (2011) Impulsive anti-periodic boundary value problem for nonlinear differential equations of fractional order. Nonlinear Anal 74:792–804
11. Ahmad B, Wang G (2011) A study of an impulsive four-point nonlocal boundary value problem of nonlinear fractional differential equations. Comput Math Appl 62:1341–1349
12. Wang J, Zhou Y, Feckan M (2012) On recent developments in the theory of boundary value problems for impulsive fractional differential equations. Comput Math Appl 64:3008–3020

A Generalized Multi-Attribute Group Decision Making with Intuitionistic Fuzzy Set

Zhifu Tao, Huayou Chen, Weiyuan Zou, Ligang Zhou and Jinpei Liu

Abstract The aim of this paper is to introduce a new decision making model called the generalized multi-attribute group decision making (GMAGDM), which provides a very general form that includes multi-attribute decision making (MADM) and multi-attribute group decision making (MAGDM) as two special cases. A GMAGDM under intuitionistic fuzzy environment is proposed. The relation between intuitionistic fuzzy set and hesitant fuzzy set is utilized. Then we propose a hesitant fuzzy TOPSIS method for the solution of the mentioned problem. Finally, a numerical example is given to illustrate the flexibility and validity of the proposed approach.

Keywords Generalized multi-attribute group decision making · MADM · MAGDM · Intuitionistic fuzzy set · Hesitant fuzzy TOPSIS

1 Introduction

Many applications involve the selection or ordering of a group of alternatives based upon their satisfaction to a collection of criteria or a number of experts, which is normally shown as a MADM problem or a MAGDM process. MAGDM

Z. Tao · H. Chen (✉) · W. Zou · L. Zhou
School of Mathematical Science, Anhui University, Hefei, Anhui 230601, China
e-mail: huayouc@126.com

Z. Tao
e-mail: zhifut_0514@163.com

W. Zou
e-mail: zwyanda123@126.com

L. Zhou
e-mail: shuiqiaozlg@126.com

J. Liu
School of Business, Anhui University, Hefei, Anhui 230601, China
e-mail: liujinpei2009@gmail.com

consists of multiple decision makers (DMs) interacting to reach a final decision, while MADM can be seen as a special condition of MAGDM where there is only one DM. For the sake of comprehensive and as commonly applied in the decision making literatures, the case of MAGDM model is considered in this paper. The case of MADM model is a straightforward simplicity of this methodology.

Dozens of studies have shown that there are several aspects to attract people's attention. They are the formats of decision information, the construction of aggregation methods and the application of the model in real problems.

Classically in MAGDM models, much of the early work focus on individual member preference as the legitimate inputs for group choice. Traditionally, the preference information provided by DMs is represented in the same format. But in practical group decision making applications, due to different cultural and educational backgrounds of DMs, it's necessary to consider different preference formats in MAGDM. Commonly used formats include: preference orderings, utility values, multiplicative preference relations, fuzzy preference relations and linguistic variables, etc. [1–6]. Totally, these types of preference information are shown with scalars. However, in real-world applications, the available information is often uncertain and fuzzy and thus, it's necessary to use another approach that is able to assess the uncertainty and vagueness in decision making. Zadeh [7] firstly introduced the fuzzy set to illustrate this fuzziness, which minimized the difference between the model and real decision process. Since then, some other forms of fuzzy decision information are proposed to illustrate different kinds of uncertainty, including the interval valued fuzzy sets (IVFs) [8], the intuitionistic fuzzy set (IFs) [9], the interval valued intuitionistic fuzzy set (IVIFs) [10], the hesitant fuzzy set (HFS) [11], etc. Fuzzy information has widely been taken into consideration to deal with imprecision of information in decision making.

In order to construct information aggregation approach for decision. In Yager [12], Yager introduced the definition of ordered weighted averaging (OWA) operator, which was contributed to cut down the negative affection of certain operation in decision making process. Related operator theories have been widely studied and applied in decision making field ([13–20]). In the meanwhile, some other ways to utilize decision information have also been developed, such as entropy/cross entropy method, programming methods, TOPSIS theory, ELECTRE I-IV, IS and TRI theory, Promethee method, etc. [19, 21–23].

When to the application of decision making theory in real-world, most crucial issues are to choose a right one from several alternatives. It can be divided as economic decision making, political decision making, etc. while combine the decision making process with actual application backgrounds [13–19, 21–27].

However, in the real process of decision making problems, a normal and special condition is that several DMs are often faced with different numbers of attributes. For example, a project needs three departments to decide whether it should be implemented, the three departments are technology department (Responsible for the assessment of technology and environment), design department (Check the economical benefits, technical feasibility and social affection of the project) and administration (Give evaluation to all properties include economic, environment,

technical and social). Then this is a special decision making problem, but it has not been researched before.

The aim of this paper is to propose the generalized multi-attribute group decision making method in which the attributes provided to DMs are not the same. We can obtain a generalization that includes normal MADM and MAGDM as its special cases. The rest of the paper is organized as follows: The following section provides some basic concepts of fuzzy set theory as preliminaries. In Sect. 3, the concept of this new type of multi-attribute group decision making process is proposed. Section 4 mainly gives the method to deal with such kind of decision making problem. The application of the proposed decision making method is addressed in Sect. 5. Finally, some conclusions are presented at the end.

2 Preliminaries

This section mainly provides with basic concepts for further discussions.

Definition 2.1 An intuitionistic fuzzy set, IFS for short, A on the universe of discourse U is defined as an object of the form $A = \{<x, u_A(x), v_A(x) > | X \in U\}$, where $u_A(x)$ and $v_A(x)$ represent the degrees of membership and non-membership of the element x to the set A, respectively, satisfying $\mathbf{0} \leq u_A(x) + v_A(x) \leq \mathbf{1}$.

As another extension and generalization of traditional fuzzy sets, hesitant fuzzy set was proposed by Torra and Narukawa [11]; it permitted the membership degree of an element to a set to be presented as several possible values between 0 and 1, which can be shown as follows:

Definition 2.2 Let X be a fixed set, a hesitant fuzzy set (HFS for short) on X is in terms of a function that when applied to X returns a subset of [0, 1], which can be represented as $E = \{<x, h_E(x) > | x \in X\}$, where $h_E(x)$ is a set of values in [0, 1], denoting the possible membership degree of the element $x \in X$ to the set E. For convince, we call $h_E(x)$ a hesitant fuzzy element (HFE).

Torra [11] showed that IFS could be transformed to be a HFS, which can be expressed as follows:

Definition 2.3 Given an IFS represented by $\{< x, u_A(x), v_A(x)\}$, the definition of corresponding HFE is straightforward: $h(x) = [u_A(x), 1 - v_A(x)]$ if $u_A(x) \neq 1 - v_A(x)$.

3 Generalized Multi-Attribute Group Decision Making with Intuitionistic Fuzzy Numbers

3.1 Concept of GMAGDM

The concept of generalized multi-attribute group decision making (GMAGDM) can be defined as follows:

Definition 3.1 Let $\{x_1, x_2, \cdots, x_m\}$ are m alternatives, $\{e_1, e_2, \cdots, e_l\}$ are l decision makers, and $A_i = \{a_1, a_2, \cdots, a_{m_i}\}$ is the corresponding attribute set of expert e_i, where $A_i \cap A_j (i \neq j, i, j \in \{1, 2, \cdots, l\})$ could be an empty set or a non-empty set and $\bigcup_{i=1}^{l} A_i = A(A = \{a_1, a_2, \cdots, a_n\}$ is attribute set, and A_i, A_j are subsets of A. Then, a generalized multi-attribute group decision making is aggregating the decision information of all alternatives under the attribute set A_i of the expert $e_i, i = 1, 2, \cdots, l$.

Note 1 The GMAGDM degenerates to be a MADM when $A_i \cap A_j = \emptyset$.

Actually, $A_i \cap A_j = \emptyset$ means that all the alternatives are judged by l decision makers under different attributes, which is the same with MADM process while thinking about the weights of decision makers.

Note 2 The GMAGDM degenerates to be a MAGDM when $A_i = A_j$.

In fact, $A_i = A_j$ illustrates that the attributes provided to each decision maker are the same, which is the same with the MAGDM problem.

Note 3 Definition 3.1 and note 1, 2 indicate that the MADM and MAGDM are two special cases of the GMAGDM.

In this paper, we mainly discuss a GMAGDM problem under intuitionistic environment.

3.2 Hesitant Fuzzy TOPSIS for GMAGDM with Intuitionistic Fuzzy Numbers

On the basis of relation between IVFs and HFs and fuzzy TOPSIS theory, we now develop a simple and straightforward method for GMAGDM with intuitionistic fuzzy decision information. The method is shown as follows:

Step 1. **Construct aggregated intuitionistic fuzzy decision matrix based on the opinions of decision makers.**

For a GMAGDM problem, let $X = \{x_1, \cdots, x_m\}$ be a finite set of alternatives, and $E = \{e_1, \cdots, e_l\}$ be the set of DMs/experts. Given attribute set of i-th DM/expert is $A_i (i = 1, 2, \cdots, l)$, where $A_i, A_j (i \neq j, i, j \in \{1, 2, \cdots, l\})$ are not required to be the same. The decision makers $e_j (j = 1, 2, \cdots, l)$ provides their judgment of alternatives under attribute set A_j, which can be described as $D_j = d_{i, \Lambda_{A_j}}$, where Λ_{A_j} is subscript set of attributes in A_j, and $d_{i, \Lambda_{A_j}}$ are intuitionistic fuzzy numbers.

Step 2. **Construct aggregated hesitant fuzzy decision matrix.**

Collect the values of alternatives under each attribute according to $D_j(j = 1, 2, \cdots, l)$; denoted as $d_{ij}^{(2)} = d_{ij}(i = 1, 2, \cdots, m; j = 1, 2, \cdots, n \in \bigcap_{k=1}^{l} \Lambda_{A_k})$, here d_{ij} is the intuitionistic value of alternative $x_i(i = 1, 2, \cdots, m)$ under attribute $a_j(j = 1, 2, \cdots, n)$. As a result, there is no less than one value of an alternative under an attribute, and the values are intuitionistic fuzzy numbers. Therefore, change $d_{ij}^{(2)}$ to be HFSs $h_{d_{ij}^{(2)}}$ based on Definition 3.1, and then the GMAGDM problem with intuitionistic fuzzy numbers comes to be a normal MADM with hesitant fuzzy attribute values.

Step 3. **Obtain hesitant fuzzy positive-ideal and negative-ideal solution.**

Extend the hesitant fuzzy numbers under the same attribute so that they have the same length and choose the positive and negative hesitant fuzzy numbers of each attribute based on Definition 2.3, denoted as

$$h^+ = \left(\max_{1 \leq i \leq m} h_{d_{i1}^{(2)}}, \max_{1 \leq i \leq m} h_{d_{i2}^{(2)}}, \cdots, \max_{1 \leq i \leq m} h_{d_{in}^{(2)}} \right), h^- = \left(\min_{1 \leq i \leq m} h_{d_{i1}^{(2)}}, \min_{1 \leq i \leq m} h_{d_{i2}^{(2)}}, \cdots, \min_{1 \leq i \leq m} h_{d_{in}^{(2)}} \right).$$

Step 4. **Calculate the separation measures.**

In order to measure separation between HFSs, Xu and Xia [28] proposed the extensions of the well-known distance measures such as Hamming distance, Euclidean distance, and Hausdorff metric under hesitant fuzzy environment, In this page, we use normalized Hamming distance [28]

$$d_{hnh}(M, N) = \frac{1}{n} \sum_{i=1}^{n} \left[\frac{1}{l_{x_i}} \sum_{j=1}^{l_{x_i}} \left| h_M^{\sigma(j)}(x_i) - h_N^{\sigma(j)}(x_i) \right| \right].$$

where h_M, h_N are two HFSs, $h_M^{\sigma(j)}, h_N^{\sigma(j)}$ are reorders of the elements in h_M, h_N.

Compute the distance between each alternative and the positive and negative hesitant fuzzy numbers, denoted as $d(x_i, h^+)$ and $d(x_i, h^-)$, respectively.

Step 5. **Calculate the relative closeness coefficient to the ideal solution.**

For each alternative, determine a ratio R equal to the distance to the negative hesitant fuzzy number divided by the sum of the distance to the negative hesitant fuzzy number and the distance to the positive hesitant fuzzy number,

$$R(x_i) = d(x_i, h^-)/(d(x_i, h^+) + d(x_i, h^-)).$$

Step 6. **Rank the alternatives.**

After the relative closeness coefficient of each alternative is determined, alternatives are ranked according to descending order of the ratio in Step 5.

4 Numerical Example

A company desires to select a new information system. Four alternatives x_i have remained in the candidate list after preliminary selection. Three departments e_j are responsible to make the final decision; they are Technology Department (TD), Evaluation Department (ED) and Administration Department (AD). Totally there are four attributes need to be considered: costs of investment (a_1); contribution to the company (a_2); effort to transform from current system (a_3); and outsourcing software developer reliability (a_4). $A_1 = \{a_1, a_3\}$, $A_2 = \{a_2, a_3, a_4\}$, $A_3 = \{a_1, a_2, a_3, a_4\}$ are three departments' attribute set, respectively. The experts evaluate the software packages with respect to their corresponding attribute sets, and construct the following fuzzy decision matrices $D_j(j = 1,2,3)$ (See Table 1).

Assume that the weights of the experts are the same because they are experts in their own fields; we utilize the former approach to select the software packages:

Step 1. Through the description of the problem, decision matrices $D_j(j = 1,2,3)$ are the corresponding intuitionistic fuzzy numbers in Table 1.

Step 2. Collect the values of alternatives, $d_{ij}^{(2)}$ can be described as:

$$d_{i1}^{(2)} = \begin{pmatrix} \langle 0.5, 0.4, 0.1 \rangle \langle 0.4, 0.5, 0.1 \rangle \\ \langle 0.4, 0.5, 0.1 \rangle \langle 0.4, 0.6, 0.0 \rangle \\ \langle 0.8, 0.2, 0.0 \rangle \langle 0.7, 0.3, 0.0 \rangle \\ \langle 0.5, 0.3, 0.2 \rangle \langle 0.7, 0.2, 0.1 \rangle \end{pmatrix},$$

$$d_{i2}^{(2)} = \begin{pmatrix} \langle 0.5, 0.5, 0.0 \rangle \langle 0.3, 0.5, 0.2 \rangle \langle 0.5, 0.4, 0.1 \rangle \\ \langle 0.6, 0.4, 0.0 \rangle \langle 0.5, 0.3, 0.2 \rangle \langle 0.7, 0.3, 0.0 \rangle \\ \langle 0.7, 0.3, 0.0 \rangle \langle 0.5, 0.2, 0.3 \rangle \langle 0.6, 0.4, 0.0 \rangle \\ \langle 0.6, 0.2, 0.2 \rangle \langle 0.7, 0.3, 0.0 \rangle \langle 0.5, 0.3, 0.2 \rangle \end{pmatrix},$$

Table 1 Intuitionistic fuzzy decision matrix (unit: 10^{-1})

	Decision matrix D_1		Decision matrix D_2			Decision matrix D_3			
	a_1	a_2	a_2	a_3	a_4	a_1	a_2	a_3	a_4
x_1	(5,4,1)	(5,5,0)	(3,5,2)	(5,2,3)	(4,5,1)	(4,5,1)	(5,4,1)	(6,2,2)	(3,5,2)
x_2	(4,5,1)	(6,4,0)	(5,3,2)	(2,6,2)	(6,4,0)	(4,6,0)	(7,3,0)	(3,5,2)	(5,5,0)
x_3	(8,2,0)	(7,3,0)	(5,2,3)	(4,4,2)	(4,6,0)	(7,3,0)	(6,4,0)	(3,5,2)	(6,2,2)
x_4	(5,3,2)	(6,2,2)	(7,3,0)	(4,2,4)	(7,1,2)	(7,2,1)	(5,3,2)	(9,1,0)	(6,4,0)

$$d_{i3}^{(2)} = \begin{pmatrix} \langle 0.5, 0.2, 0.3 \rangle \langle 0.6, 0.2, 0.2 \rangle \\ \langle 0.2, 0.6, 0.2 \rangle \langle 0.3, 0.5, 0.2 \rangle \\ \langle 0.4, 0.4, 0.2 \rangle \langle 0.3, 0.5, 0.2 \rangle \\ \langle 0.4, 0.2, 0.4 \rangle \langle 0.9, 0.1, 0.0 \rangle \end{pmatrix},$$

$$d_{i4}^{(2)} = \begin{pmatrix} \langle 0.4, 0.5, 0.1 \rangle \langle 0.3, 0.5, 0.2 \rangle \\ \langle 0.6, 0.4, 0.0 \rangle \langle 0.5, 0.5, 0.0 \rangle \\ \langle 0.4, 0.6, 0.0 \rangle \langle 0.6, 0.2, 0.2 \rangle \\ \langle 0.7, 0.1, 0.2 \rangle \langle 0.6, 0.4, 0.0 \rangle \end{pmatrix}$$

Then change them to be hesitant fuzzy values and shows as follows:

$$h_{d_{i1}^{(2)}} = \begin{pmatrix} (0.4, 0.5, 0.6) \\ (0.4, 0.5) \\ (0.7, 0.8) \\ (0.5, 0.7, 0.8) \end{pmatrix}, \quad h_{d_{i2}^{(2)}} = \begin{pmatrix} (0.3, 0.5, 0.6) \\ (0.5, 0.6, 0.7) \\ (0.5, 0.6, 0.7, 0.8) \\ (0.5, 0.6, 0.7, 0.8) \end{pmatrix},$$

$$h_{d_{i3}^{(2)}} = \begin{pmatrix} (0.5, 0.6, 0.8) \\ (0.2, 0.3, 0.4, 0.5) \\ (0.3, 0.4, 0.5, 0.6) \\ (0.4, 0.8, 0.9) \end{pmatrix}, \quad h_{d_{i4}^{(2)}} = \begin{pmatrix} (0.3, 0.4, 0.5) \\ (0.5, 0.6) \\ (0.4, 0.6, 0.8) \\ (0.6, 0.7, 0.9) \end{pmatrix}.$$

Step 3. Extend hesitant fuzzy numbers under each attribute and indentify the positive and negative hesitant fuzzy set, denoted as follows:

$d^+ = ((0.7,0.8,0.8),(0.5,0.6,0.7,0.8),(0.4,0.8,0.9,0.9),(0.6,0.7,0.9));$
$d^- = ((0.4,0.5,0.5),(0.3,0.5,0.6,0.6),(0.2,0.3,0.4,0.5),(0.3,0.4,0.5)).$

Step 4. Compute the distance between each alternative and separation measures.

$d(x_1,h^-) = 0.0896$, $d(x_1,h^+) = 0.2063$; $d(x_2,h^-) = 0.0729$, $d(x_2,h^+) = 0.2229$;
$d(x_3,h^-) = 0.1875$, $d(x_3,h^+) = 0.1083$; $d(x_4,h^-) = 0.2708$, $d(x_4,h^+) = 0.2958$.

Step 5. For each alternative, indentify the ratio R, and the result is,

$R(x_1) = 0.3028$, $R(x_2) = 0.2465$, $R(x_3) = 0.6338$, $R(x_4) = 0.9155$.

Step 6. Rank the software according to the result in Step5, the order should be:
$x_4 \succ x_3 \succ x_1 \succ x_2$.

where "$\succ$" denotes "be superior to". Therefore, x_4 is the best alternative.

By comparison with the traditional information aggregation approach, the proposed GMAGDM model presents some advantages:

- A lot of computation among intuitionistic fuzzy numbers and possible loss of information are avoided.
- The application of hesitant fuzzy set is established, through which one can also found the relation of two fuzzy sets in real, but not just stay in the written.

5 Conclusions

In this paper, we introduced the GMAGDM model, which can be regarded as a generalization of normal MADM and MAGDM where attribute sets provided for DMs are not required to be the same. The characteristic of IVFs and HFs were utilized, and the hesitant fuzzy TOPSIS theory was introduced and applied in the solving of GMAGDM. Numerical example illustrated that this method had some advantages and we also proposed a new type of group decision making.

References

1. Tanino T (1984) Fuzzy preference orderings in group decision making. Fuzzy Sets Syst, 12:117–131
2. Delgado M, Herrera F, Herrera-Viedma E, Martinez L (1998) Combining numerical and linguistic information in group decision making. Inf Sci 107:177–194
3. Herrera F, Herrera-Viedma E, Chiclana F (2001) Multiperson decision-making based on multiplicative preference relations. Eur J Oper Res 129:372–385
4. Kameda T, Tindale RS, Davis JH (2002) Cognitions, preferences, and social sharedness: past, present, and future directions in group decision making. In: Schneider SL, Shanteau J (eds) Emerging perspectives on judgment and decision research. Cambridge University Press, Cambridge
5. Chen HY, Zhou LG, Han B (2011) On compatibility of uncertain additive linguistic preference relations and its application in the group decision making. Knowl-Based Syst 24:816–823
6. Merigo JM, Gil-Lafuente AM, Zhou LG, Chen HY (2011) A generalization of the linguistic aggregation operator and its application in decision making. J Syst Eng Electron 22:1–5
7. Zadeh LA (1965) Fuzzy sets, information and control. 8:338–353
8. Zadeh LA (1975) The concept of a linguistic and application to approximate reasoning I. Inf Sci, 8:199–249
9. Krassimir T (1988) Atanassov: intuitionistic fuzzy sets. Fuzzy Sets Syst 20:87–90
10. Atanassov K, Gargov G (1989) Interval valued intuitionistic fuzzy sets. Fuzzy Sets Syst, 31:343–349
11. Torra V, Narukawa Y (2009) On hesitant fuzzy sets and decision, FUZZ-IEEE, Korea, Aug pp 20–24
12. Yager R (1993) Families of OWA operators. Fuzzy Sets Systems. 59:125–148
13. Zhou LG, Chen HY (2012) A generalization of the power aggregation operators for linguistic environment and its application in group decision making. Knowl Based Syst, 26:216–224
14. Chen HY, Zhou LG (2012) A relative entropy approach to group decision making with interval reciprocal relations based on COWA Operator. Group Decis Negot. 21:585–599
15. Zhou LG, Chen HY (2011) Continuous generalized OWA operator and its application to decision making. Fuzzy Sets Syst, 168:18–34
16. Zhou LG, Chen HY, Liu JP (2012) Generalized power aggregation operators and their applications in group decision making. Comput Ind Eng 62:989–999

17. Zhou LG, Chen HY, Liu JP (2012) Continuous ordered weighted distance measure and its application to multiple attribute group decision making. Group Decis Negot. doi:10.1007/s10726-012-9289-3
18. Zhou LG, Chen HY, Merigó JM, Gil-Lafuente AM (2012) Uncertain generalized aggregation operators. Expert Syst Appl 39:1105–1117
19. Chen HY, Zhou LG (2011) An approach to group decision making with interval fuzzy preference relations based on induced generalized continuous ordered weighted averaging operator. Expert Syst Appl 38:13432–13440
20. Liu PD (2013) Some geometric aggregation operators based on interval intuitionistic uncertain linguistic variables and their application to group decision making. Appl Math Model, 37:2430–2444
21. Xia MM, Xu ZS (2012) Entropy/cross entropy-based group decision making under intuitionistic fuzzy environment. Inf Fusion 13:31–47
22. Wan ZJ, Li KW, Xu JH (2011) A mathematical programming approach to multi-attribute decision making with interval-valued intuitionistic fuzzy assessment information. Expert Systems with Applications, 38:12462–12469
23. De Smet Y, Mareschal B, Verly C (2009) Extending the PROMETHEE II method to continuous and combinatorial multi-objective optimization problems: a first model. In: Proceedings of the 2009 IEEE IEEM. 1608–1611
24. Zhang SF, Liu SY (2011) A GRA-based intuitionistic fuzzy multi-criteria group decision making method for personnel selection. Expert Syst Appl 38:11401–11405
25. Vahdani B, Mousavi SM, Tavakkoli-Moghaddam R, Hashemi H (2013) A new design of the elimination and choice translating reality method for multi-criteria group decision-making in an intuitionistic fuzzy environment. Appl Math Model 37:1781–1799
26. Chiclana F, Tapia García JM, del Moral MJ, Herrera-Viedma E (2013) A statistical comparative study of different similarity measures of consensus in group decision making. Inf Sci 221:110–123
27. Palomares I, Rodríguez RM, Martínez L (2013) An attitude-driven web consensus support system for heterogeneous group decision making. Expert Syst Appl 40:139–149
28. Xu ZS, Xia MM (2011) On distance and correlation measures of hesitant fuzzy information. Int J Intell Syst 26:410–425

Asymptotic Optimality of Balanced Routing in a Multiclass G/G/1 Queue

Xiao Xiao and Lu Wu

Abstract In this paper we consider a multiclass G/G/1 queueing system. When the server is idle, we use a balanced routing control policy to determine what kind of customer is served. Under such a balanced policy, we derive the diffusion limits of the queue length processes and the workload processes. The diffusion limits are the same for these processes regardless the choice of c as long as $c \geq 2$. We further show that the proposed balanced routing policy for any fixed $c \geq 2$ is asymptotically optimal in the sense that it minimizes the workload over all time in the diffusion limit.

Keywords G/G/1 queueing system · Balanced routing · Asymptotic optimality

1 Introduction

We consider a multiclass G/G/1 queueing system. When the server is idle, we use a balanced routing control policy to determine what kind of customer is served. This problem arises in a number of important applications. For example, it can be viewed as a order model of Van Mieghem [1]. It is interesting to think of our system as modeling order fulfillment at a firm which dynamically receives orders from customers for several different types or classes of goods and services. In order not to keep one kind of customer's orders in the firm for too long, we need to

X. Xiao (✉)
Department of Mathematics, Anhui University of Science and Technology, Huainan
232007, People's Republic of China
e-mail: kerryshaw@sjtu.edu.cn

L. Wu
Department of Economics, Anhui University of Science and Technology, Huainan 232007,
People's Republic of China
e-mail: wulu08cn@gmail.com

settle the problem of routing control in such a system. The purpose of this paper is to show how the firm should sequence the different orders that are competing for its scarce resources.

Load balancing is a canonical method for sharing resources among different customers efficiently. Two versions of the load balancing problem have been studied in the literature: static and dynamic. The static version was first analyzed by Azar et al. [2] using the balls-and-bins model. In this paper, we are interested in the dynamic version, in Vvednskaya et al. [3], a routing policy that is based on the partial queue length information is proposed. The model was also studied by Mitzenmacher [4].

Chen and Ye [5] extends and complements the previous works in several ways. Firstly, based on Mitzenmacher's, they establish the asymptotic optimality of the balanced routing under more general queueing systems. Secondly, they complements the work of He and Down [6], they rigorously establish the heavy traffic limit theorem for the queueing system under balance routing. Compared to Chen and Ye [5], the distinguishing feature of our analysis is that here we consider a multiclass G/G/1 queueing system. We extend the balanced routing control policy to multiclass G/G/1 queueing system.

The paper is organized as follows. In the next section we present our model. Section 3 analyze the fluid limit of the queueing processes, and establish the uniform attraction property of the fluid limit. In Sect. 4, we show that for any $c \geq 2$, the balanced routing control is asymptotically optimal. We conclude in Sect. 5 with extensions and discussion.

2 The Model

We consider a general single-server multiclass queueing system. Customers are categorized into L different classes depending on their specific arrival patterns. Class $l \in \mathcal{L} = \{1, \dots, L\}$ customers arrive at the system following a renewal process with arrival rate λ_l. At the server, the customers are served at a rate of μ_l according to the first-in first-out (FIFO) discipline.

For class l customers, denote the interarrival times between consecutive customers as $u_{l,k}$, $l \in \mathcal{L}$; $k \geq 1$, and denote the service time of kth customer at the server by $v_{l,k}$, $l \in \mathcal{L}$; $k \geq 1$. We assume that $\{u_{l,k}, k \geq 1\}$, $l \in \mathcal{L}$, and $\{v_{l,k}, k \geq 1\}$, $l \in \mathcal{L}$, are i.i.d. random sequence, all with finite second moments. In particular, $u_{l,k}$ are random variables with mean $1/\lambda_l$ and variance a_l^2, $l \in \mathcal{L}$; and $v_{l,k}$ are random variables with mean $1/\mu_l$ and variance b_l^2, $l \in \mathcal{L}$. Denote the traffic intensity as $\rho_l = \lambda_l/\mu_l$, $l \in \mathcal{L}$. For any $\mathcal{L}' \subseteq \mathcal{L}$, let $\lambda_{\mathcal{L}'} = \sum_{l \in \mathcal{L}'} \lambda_l$ and $\mu_{\mathcal{L}'} = \sum_{l \in \mathcal{L}'} \mu_l$. Let $\lambda_0 = \lambda_{\mathcal{L}}/L$ denote the average arrival rate and $\mu_0 = \mu_{\mathcal{L}}/L$ the average service rate.

We call $A(t) = (A_l(t))_{l \in \mathcal{L}}$ the arrival process, where $A_l(t)$ represents the number of class l customers that have arrived during the time interval $[0, t]$; and call $S(t) = (S_l(t))_{l \in \mathcal{L}}$ the service process, where $S_l(t)$ is the number of class l customers that are served during the first t time units that the server devotes to class l.

$$U_l(0) = 0, \quad U_l(k) = \sum_{k'=1}^{k} u_{l,k'}, \quad k \geq 1, \; l \in \mathcal{L},$$

$$V_l(0) = 0, \quad V_l(k) = \sum_{k'=1}^{k} v_{l,k'}, \quad k \geq 1, \; l \in \mathcal{L},$$

$$A_l(t) = \sup\{k \geq 0 : U_l(k) \leq t\}, \text{ and } S_l(t) = \sup\{k \geq 0 : V_l(k) \leq t\}.$$

Let $q = (q_1, \ldots, q_L)'$ be a nonnegative integer valued vector (with q_l denote the queue length of class l customers). Let $\pi(q) = (j_1, j_2, \ldots, j_L)'$ be a permutation of indices $\mathcal{L}$ such that $q_{j_1} \geq q_{j_2} \geq \cdots \geq q_{j_L}$, and $\pi_l(q)$ be the set of the first l components of $\pi(q)$. Let $\phi(j, k; \pi)$ be a zero–one random variable, and $\phi(j, k; \pi) = 1$ represents that the kth service is dispatched to the class j customers. For $\forall q$ and $\forall \mathcal{L}' \subseteq \mathcal{L}$ that satisfy $\pi_l(q) = \mathcal{L}'$ for some l $(1 \leq l \leq L)$,

$$\Phi(\mathcal{L}', k) := \sum_{j \in \mathcal{L}'} \phi(j, k; \pi(q)). \tag{1}$$

We assume that the routing sequence is independent of the arrival process and the service time process. For ease of analysis, we introduce

$$R(\mathcal{L}', 0) = 0, \; R(\mathcal{L}', k) = \sum_{k'=1}^{k} \Phi(\mathcal{L}', k'), \; k \geq 1, \text{ for all } \mathcal{L}' \subseteq \mathcal{L}.$$

3 Fluid Limit and Uniform Attraction

Here we start with presenting the main processes in this paper. First, let $Q(t) = \{Q_l(t)\}_{l \in \mathcal{L}}$ be the queue length at time t, where $Q_l(t)$ denote the total number of customers of class l in the system at time t. Then, $B_l(t)$ is the number of customers of class l that are served during $[0, t]$, $T_l(t)$ represents total amount of the time during $[0, t]$ that the server allocates to class l. The above can be described as follows:

$$B_l(t) = \sum_{k=1}^{S(t)} \phi(l, k; \pi(Q(V(k)-))), \; \sum_{l \in \mathcal{L}} B_l(t) = S(t) = \sum_{l \in \mathcal{L}} S_l(t), B_l(t) = S_l(T_l(t)).$$

$$\tag{2}$$

The derived process that characterizes the dynamics of the queueing system is the queue-length process:

$$Q_l(t) = Q_l(0) + A_l(t) - S_l(T_l(t)) \geq 0, \tag{3}$$

$Q_l(0)$ is the initial number of customers of class l in the system, $l \in \mathcal{L}$.

Now define the workload process and the idling process: $W(t) = (W_l(t))_{l \in \mathcal{L}}$ and $Y(t) = (Y_l(t))_{l \in \mathcal{L}}$, where

$$W_l(t) = V_l(Q_l(0)) + V_l(A_l(t)) - T_l(t) = V_l(Q_l(t)) = \frac{1}{\mu_l} Q_l(t), \tag{4}$$

and $Y_l(t) = t - \sum_{l \in \mathcal{L}} T_l(t)$. In addition, we define $W_0(t)$ and $Y_0(t)$ as follows,

$$W_0(t) = \frac{1}{\mu_{\mathcal{L}}} \sum_{l \in \mathcal{L}} \mu_l W_l(t) = \frac{1}{\mu_{\mathcal{L}}} \sum_{l \in \mathcal{L}} Q_l(t), \tag{5}$$

$$Y_0(t) = \frac{1}{\mu_{\mathcal{L}}} \sum_{l \in \mathcal{L}} \mu_l Y_l(t) = t - \frac{1}{\mu_{\mathcal{L}}} \sum_{l \in \mathcal{L}} \mu_l T_l(t). \tag{6}$$

Theorem 1 (Fluid Limit) *Let M be a given positive constant, and suppose $|\bar{Q}^n(0)| = \sum_{l \in \mathcal{L}} \bar{Q}_l^n(0) \leq M$ for all n. Then, for any subsequence of fluid scaled processes, there exists a further subsequence, denoted by $\mathcal{N}$, such that, along $\mathcal{N}$,*

$$\begin{aligned}
\left(\bar{Q}^n(t), \ \bar{B}^n(t), \ \bar{T}^n(t), \ \bar{W}^n(t), \ \bar{W}_0^n(t), \ \bar{Y}^n(t), \ \bar{Y}_0^n(t) \right) \rightarrow \\
\left(\bar{Q}(t), \ \bar{B}(t), \ \bar{T}(t), \ \bar{W}(t), \ \bar{W}_0(t), \ \bar{Y}(t), \ \bar{Y}_0(t) \right), \ u.o.c
\end{aligned} \tag{7}$$

Theorem 2 (Uniform Attraction) *Let M be a given positive constant, and suppose $|\bar{Q}^n(0)| = \sum_{l \in \mathcal{L}} \bar{Q}_l^n(0) \leq M$ for all n. Assume that the heavy traffic condition $\rho = 1$ holds. Then, there exists a time $T_M > 0$ such that, all fluid levels are the same after the time T_M and the fluid levels remain fixed afterward: for all $l \in \mathcal{L}$,*

$$\bar{Q}_l(t) = \mu_0 \bar{W}_0(t) (= \mu_0 \bar{W}_0(0)), \ \text{for } t \geq T_M. \tag{8}$$

4 Diffusion Limit and Asymptotic Optimality

We apply the standard diffusion scaling (along with centering) to the derived process:

$$\hat{Q}^n(t) := \frac{1}{n} Q^n(n^2 t), \ \hat{W}^n(t) := \frac{1}{n} W^n(n^2 t), \ \hat{W}_0^n(t) := \frac{1}{n} W_0^n(n^2 t). \tag{9}$$

We can rewrite the unscaled system workload process for the nth network, and apply the diffusion scaling to the equation, we have

$$\hat{W}_0^n(t) = \frac{1}{n} W_0^n(n^2 t) = \hat{X}_0^n(t) + \hat{Y}_0^n(t), \tag{10}$$

where

$$\hat{X}_0^n(t) := \frac{1}{n\mu_{\mathcal{L}}^n} \sum_{l \in \mathcal{L}} \mu_l^n V_l^n(\tilde{Q}_l^n(0))$$

$$+ \frac{1}{\mu_{\mathcal{L}}^n} \sum_{l \in \mathcal{L}} \mu_l^n V_l^n(\hat{A}_l^n(t) - \hat{S}^n(\tilde{T}_l^n(t))) + n(\rho^n - 1)t, \tag{11}$$

$$\hat{Y}_0^n(t) := n\left(t - \frac{1}{\mu_{\mathcal{L}}^n} \sum_{l \in \mathcal{L}} \mu_l^n \tilde{T}_l^n(t)\right), \tag{12}$$

For each n, we also have

$$\hat{Y}_0^n(t) \text{ is nondecreasing in } t \geq 0, \text{ and } \hat{Y}_0^n(0) = 0. \tag{13}$$

In view of (4) and with the centering procedure as in the above, we can write the following equality for the workload process,

$$\hat{W}_l^n(t) = \frac{1}{\mu_l^n} \hat{Q}_l^n(t) + \hat{V}_l^n(\tilde{Q}_l^n(0)) + \hat{V}_l^n(\tilde{A}_l^n(t)) + \frac{1}{\mu_l^n} \hat{S}_l^n(\tilde{B}_l^n(t)). \tag{14}$$

By the functional central limit theorem for the renewal process (refer to Chen and Yao [7], Chap. 5), we have

$$(\hat{A}^n, \hat{V}^n, \hat{S}^n) \Rightarrow (\hat{A}, \hat{V}, \hat{S}), \text{as } n \to \infty, \tag{15}$$

For ease of exposition, assume that

$$Q_l^n(0) = 0, \text{ for all } l \in \mathcal{L} \text{ and all } n. \tag{16}$$

It follows from Theorem 1 that as $n \to \infty$,

$$(\tilde{Q}_l^n(0), \tilde{T}_l^n(t)) \to (0, \rho_l t), \text{ u.o.c}, \tag{17}$$

It follows from the random time-change theorem (refer to Chen and Yao [7], Chap. 5), the process $\hat{X}^n(t)$ converges weakly as follows,

$$\hat{X}_0^n(t) \Rightarrow \hat{X}_0(t) := \theta t + \sum_{l \in \mathcal{L}} V_k(\hat{A}_l(t) - \hat{S}(\rho_l t)), \text{ as } n \to \infty, \tag{18}$$

where the limit $\bar{X}_0 = \{\hat{X}_0(t), t \geq 0\}$ is a Brownian motion.

We show that the derived diffusion scaled process approach to some limits described by $\hat{W}_0(t)$, $\hat{Y}_0(t)$ in the next theorem. The limiting process are characterized by the following relations for all $t \geq 0$ and $l \in \mathcal{L}$:

$$\hat{W}_0(t) = \hat{X}_0(t) + \hat{Y}_0(t) \geq 0; \tag{19}$$

$$\hat{Y}_0(t) \text{ is nondecreasing in } t, \text{ with } \hat{Y}_0(0) = 0; \tag{20}$$

$$\int_0^\infty \hat{W}_0(t) d\hat{Y}_0(t) = 0. \tag{21}$$

Theorem 3 *Suppose that the heavy traffic condition $\rho = 1$ holds. Under the balanced routing policy, we have the following results.*

(a) **Diffusion Limit:** *The following weak convergence holds when $n \to \infty$:*

$$\big(\hat{W}_0^n(t), \ \hat{Y}_0^n(t), \ \hat{Q}^n(t)\big) \Rightarrow \big(\hat{W}_0(t), \ \hat{Y}_0(t), \ \hat{Q}(t)\big),$$

in particular, $\hat{W}_0$ is a single-dimensional RBM, $\hat{Y}_0$ is the associated regulator, and $\hat{Q}(t) = Q^*\big(\hat{W}_0(t)\big)$, is the fixed point with all components equal to $\mu_0 \hat{W}_0(t)$.

(b) **Asymptotic optimality:** The balanced routing is asymptotically optimal in the following sense: Let $\hat{W}_0^{n,G}$ and $\hat{Q}^{n,G}$ denote the scaled workload and queue length process associated with any feasible routing scheme $G.$. Then, for all $t \geq 0$ and $u \geq 0$, we have

$$\liminf_{n \to \infty} P\Big\{\hat{W}_0^{n,G}(t) > u\Big\} \geq P\big\{\hat{W}_0(t) > u\big\}, \text{ and} \tag{22}$$

$$\liminf_{n \to \infty} P\Big\{\max_l \hat{Q}_l^{n,G}(t) > u\Big\} \geq P\Big\{\max_l \hat{Q}_l(t) > u\Big\}. \tag{23}$$

5 Concluding Remarks

In this paper, we extend the balanced routing control policy to multiclass G/G/1 queueing system. We prove that, under the diffusion scale to minimize the total workload process and the maximum queue length process, for any $c \geq 2$, the balanced routing control is asymptotically optimal. In heavy traffic, the multi-server simplifies to a single server with service capacity equal to the sum of the parallel servers, and the analysis still applies.

References

1. Stolyar AL (2001) Max-weight scheduling in a generalized switch: state space collapse and workload minimization in heavy traffic. Ann Appl Probab 14x:1–53
2. Azar Y, Broder A, Karlin A, Upfal E (1999) Balanced allocations. SIAM J Comput 29:1180–1200

3. Vvednskaya ND, Dobrushin RL, Karpelevich FI (1996) Queueing system with selection of the shortest of two queues: an asymptotic approach. Probl Inf Transm 32:15–27
4. Mitzenmacher M (2001) The power of two choices in randomized load balancing. IEEE Trans Parallel Distrib Comput 12:1094–1104
5. Chen H, Ye HQ (2012) Asymptotic optimality of balanced routing. Oper Res 60(1):163–179
6. He YT and Down DG (2008) Limited choice and locality considerations for load balancing. Perform Eval 65:670–687
7. Chen H, Yao DD (2001) Fundamentals of Queueing networks: performance, asymptotics, and optimization. Springer, New York

The Optimization of DNA Encodings Based on GAFSA/GA Algorithm

Juan Hu, Dong Li, Li-li Zhang and Zhixiang Yin

Abstract The design of DNA sequence is important in improving the reliability of DNA computing. Some appropriate constrained terms that DNA sequence should satisfy are selected, and then the evaluation formulas of each DNA individual corresponding to the selected constrained terms are proposed. The paper analyzes the objective and several constraints of DNA encoding, it builds a combinational optimization model. A Global Artificial Fish Swarm algorithm/ Genetic Algorithm (GAFSA/GA) is proposed to produce DNA encoding sequences. The result shows that the DNA sequences produced by GAFSA/GA have better quality than that produce by the genetic algorithm.

Keywords DNA computing · DNA encoding · Combinational optimization · Artificial fish swarm

J. Hu (✉)
The Foundation Department of Huainan Vocational Technical College,
Huainan 232001, China
e-mail: hujuan64072149@163.com

D. Li
The Electronics Set of Huainan Industrial School, Huainan 232001, China

L. Zhang · Z. Yin
Department of Mathematics and Physics, Anhui University of Science
and Technology, Huainan 232001, China

1 Introduction

In 1994, Adleman used the DNA to solve the toward way of Hamilton road problem, this expanded a sexual achievement, Which carrying on a calculating modern era through a bio-chemical reaction with the DNA member. From now on, The model of DNA computer has highly and abreast compute with it, the sea quantity is saving ability at polynomial time inside worked out a complete problem of a lot of NPs, also people have tremendous interesting on DNA calculator.

Genetic algorithm (GA) is first proposed of Holland at the beginning of the nineteen seventies, which is based on natural selection and genetic theory, and it is the optimization search method which combining the fittest rules and populations within chromosomes random information exchange system in the biological evolution process of survival [1]. Because of its robustness, especially suitable for the processing of complex, nonlinear problem, which is based on DNA GA was used to search the better coding sequence [2]. Wood and Chen proposed by GA a sequence design and applied to solve the optimal routing problem [3]. But the genetic algorithm shortcomings are easy to generate inbreeding coefficient, prematurity, thereby falling into local extremum.

In this paper, based on the traditional genetic algorithms to search blindly shortcomings, combined with global artificial fish-swarm algorithm is simple and easy to realize, has achieved global extremum, avoid falling into local extremum good ability to search space, also have a certain adaptive ability, and the algorithm is not sensitive to initial value and parameter, has strong robustness and good convergence characteristics. Proposed one kind based on the genetic algorithm and the global artificial fish swarm algorithm hybrid optimization method (GAFSA/ GA) for DNA coding problem, thus overcoming the genetic evolution of blindness, enhancing evolutionary efficiency, having the optimize DNA code.

2 DNA Coding Sequence Design and Constraint Analysis

2.1 DNA Coding Problem

Because of the DNA computation is through the DNA molecular hybridization to complete, so the coding problem research is to maximize the hope is encoded by the DNA molecules are able to complete hybridization, and do not want to appear incomplete complementary DNA molecular hybridization and complete complementary DNA molecules do not cross the phenomenon, that can improving computational efficiency and DNA reliability.

The DNA calculation of the coding problem usually described as: one composed of four letters collection $\sum = \{A, T, C, G\}$, if DNA molecular length is n, the code set is S, obviously $|S| = 4^n$. Seek $C \subseteq S$, make $\forall x_i, x_j \in C \, \tau(x_i, x_j) \geq k$, of which $k \in Z^+, \tau$ is the standard for evaluating. It is Clear that the evaluation

criteria τ is more string enter, the options for coded number $|C|$ is less. Obviously, in the coding of coding quality and quantity is concern. Because the higher the quality, the reliability is better; the greater the number, scope of application is enlarged. However, in reality they are contradictory. This requires to maximize the obtained coding set in the premise of quality assurance.

2.2 Fitness Function

This article defines the optimization problem as the maximum value problem, using the weighted average method to process each individual DNA the constraint evaluation function [4].

$$Fitness(x_i) = \sum_{j=1}^{m} \omega_j f_j$$

In which m is the number of constraints, ω_j for each constraint term f_j weighting, and have

$$f_j \in \left\{ f_{H-measure}(x_i), f_{Similarity}(x_i), f_{Con}(x_i), f_{Hairpin}(x_i), f_{GC}(x_i), f_{Tm}(x_i) \right\}$$

3 GAFSA/GA

3.1 The Basic Genetic Algorithm

Genetic algorithm is a calculation model simulating the Darwin biological theory of evolution by natural selection and genetic mechanism of the process of biological evolution, is the optimal solution search method simulating the natural evolutionary process, it initiated by the United States Michigan university professor Holland in 1975. Genetic algorithm with decision variable encoding as the operation object, without the need to solve the problem of domain specific information, at the same time it is not constricted by the search space continuous or differentiable. Genetic algorithm can be divided into 4 main sections: the initial state of the establishment, fitness function, genetic operation and setting control parameters.

3.2 GAFSA

3.2.1 Artificial Fish Swarm Behavior Description

In a body of water, the fish often can automatically or trailing other fish to find the more nutrients places, so where the number of fish living most is the most of the nutrients in general, artificial fish swarm algorithm is based on this characteristic, through the construction of artificial fish to imitate fish foraging, gathering and rear-end behavior, thus realizing the optimization, the following are some typical behavior of fish: (1) foraging behavior: in general the fish in the water is randomly free swimming, when found the food, will be to increase gradually in the direction of the fast food tour. (2) gathering behavior: fish moving process in order to guarantee their survival and avoid hazards naturally clustered, fish party obey three rules: try to avoid close partner too crowded; try and close the mean direction of partners; try to close partner center mobile. (3) following behavior: when one or several fish find the food, near partners will follow the rapid arrival of food.

3.2.2 Mathematical Model of Artificial Fish

Hypothesis in an n-dimensional object search space, a n consists of a group of artificial fish, each individual state of artificial fish can be expressed as $X = (x_1, x_2, \cdots, x_n)$, which $x_i (i = 1, 2, \cdots, n)$ is to find optimal variables; artificial fish of the current location of food concentration expressed as $Y = f(x)$, where Y as the objective function value; artificial fish distances among individuals expressed as $d_{i,j} = \|X_i - X_j\|$; *Visual* as artificial fish perceptual range; *Step* as the artificial fish moving step; δ as the congestion factor. In algorithm artificial fish, i is the artificial fish school, the neighborhood is defined as $N = \{X_j | d_{i,j} < Visual\}$, At global version of artificial fish swarm algorithm artificial fish i in the neighborhood defined as the entire fish, the artificial fish perform cluster, rear-end behavior, each artificial fish and fish of all other artificial fish "communication", all of the artificial fish sharing the information of the optimal industrial fish and the information of the entire fish center position, speed up information transmission speed, improved algorithm convergence speed and running speed.

1. Random behavior: (AF-Random) refers to the artificial fish in view of random movement, when found in food, would be to increase gradually in the direction of the fast moving food.
2. Foraging behavior: In their field of vision artificial fish according to (1) randomly select a state X_j, calculate for the objective function values and compare, if found Y_j outgo Y_i, then the artificial fish move directly to X_j; otherwise, in the view of random X_i according to (1) randomly select a state X_j determining whether they meet forward conditions; repeated a few times, if still not satisfied

with forward condition, according to the (2) randomly moves one step and sending X_i to reach a new state.

$$X_j = X_i + (2rand() - 1)Visual \qquad (1)$$

$$X_i = X_i(1 + (2rand() - 1)Step) \qquad (2)$$

3. Gathering behavior: artificial fish X_i search in its field partners number M and center position X_c, if $Y_c M > \delta Y_i$, show the partner center position better and less crowded, according to (3), X_i moves one step towards the central position in partner, otherwise the implementation of foraging behavior.

$$X_i = X_i + rand()Step \frac{X_c - X_i}{\|X_c - X_i\|} \qquad (3)$$

4. Following behavior: artificial fish X_i search the optimal partner X_{gbest} in its vision of all partners in the function, if $Y_{gbest} M > \delta Y_i$, show that the optimal partner around not too crowded, according to (4), X_i moves one step towards this partner, otherwise the implementation of foraging behavior.

$$X_i = X_i + rand()Step \frac{X_{gbest} - X_i}{\|X_{gbest} - X_i\|} \qquad (4)$$

5. Artificial fish vision and step dynamic adjustment [4]:

$$\begin{cases} Visual = Visual \times a + Visual_{min} \\ \quad Step = Step \times a + Step_{min} \\ \quad a = \exp(-30 \times (t/T_{max})^s) \end{cases} \qquad (5)$$

The artificial fish swarm algorithm steps are as follows:

Step1: Determination the fish scale N, randomly generated N artificial fish body, consisting of the initial population, at the same time setting related parameters.

Step2: Calculate the initial fish each individual fitness, and the optimal working state of fish and its value to the bulletin board.

Step3: Individual through foraging, gathering, rear-end behavior update themselves, generate new fish.

Step4: Evaluation of all individual, choose the best individual and bulletin boards were recorded for comparison, if better, with their status updates. On board recording

Step5: According to formula (5) and the artificial fish vision and step size dynamically

Step6: Judging whether a termination condition is satisfied, if satisfied, then the output board records, algorithm terminates; if it is not satisfied, then executing Step 3.

3.3 GAFSA/GA

The algorithm is based on genetic algorithm foundation, at the same time the artificial fish swarm algorithm genetic algorithm as an operator, a specific algorithm steps are as follows:

1. parameter set, and a randomly generated initial population;
2. calculate for each individual fitness function value, and according to fitness function value sequence [5];
3. judges whether meet the objective conditions (including program convergence and reach the designated the evolution algebra), if satisfied, the end of the process, output the results; otherwise proceed to next step;
4. update individual populations, according to the fitness function value to determine the part of the individual directly into the population of the next generation, the remaining individual through the GAFSA/GA algorithm to enter after the population of the next generation;
5. pairs of new generation implementing genetic algorithm copy, crossover and mutation operation, to step 2.

4 Simulation Results and Analysis

In view of the above constraint condition and the object function design coding sequence model, in the Matlab7.0 environment, using GAFSA/GA algorithm simulation, the operating environment is Pentium Dual E2104, 1.6 GHz, 512 MB, Microsoft XP. Parameters are set as follows [5]: (1) the basic genetic algorithm evolution: maximum algebra of 300, the population size of 20, DNA coding sequence of length 20, crossover rate of 0.85, mutation rate of 0.005. (2): GAFSA algorithm artificial fish swarm size M = 20 repeated attempts, try_number = 5, s = 2, Visualmin = 0.001, Stepmin = 0.0002. Evolution algebra 100, continuous running 50 times.

   In this paper, the global artificial/genetic sequence design algorithm to generate DNA sequence and Arita and others in the literature [6] genetic algorithm coding sequences were compared, and gives 7 more suitable for Adleman to Hamilton road problem of DNA sequence. And the same coding evaluation standards were analyzed, compared the results as shown in Table 1 as shown.

   As you can see from Fig. 1, the GASFA/GA algorithm to generate the DNA coding sequence in the hairpin structure, the content of GC is superior to GA algorithm to generate the DNA coding sequence, can see GASFA/GA sequence melting temperature Tm ∈ [56.2037, 57.7121], GA algorithm produces a sequence of melting temperature of Tm ∈ [48.445 1, 69.200 7], GAFSA/GA generated sequence melting temperature change is smaller, this means that the GAFSA/GA algorithm has low does not exactly match the probability of generating a double chain

   GASFA/GA algorithm to solve the problem of DNA coding show good effect, this is because it follow the current search to optimal value to find the global optimum, greatly reduces the computational complexity. At the same time groups of artificial fish parallel computing has a faster convergence rate than the genetic algorithm average only about 50 iterations to obtain quality DNA coding sequence set, this often requires thousands of iterative genetic algorithm, is a big breakthrough.

   To sum up, this algorithm generates code has more stable thermodynamic stability

Table 1 The compared results of GASFA/GA and GA

DNA coding sequence	H	C	S	Hairpin	Tm	GC %
GAFSA/GA						
GCTGTGCACATCTGACTATG	54	0	45	0	56.5746	50
TGTACATCTCTGACCCTGCT	54	9	51	0	56.2037	50
CGATGTTCGCAGTGCTCACT	57	0	37	0	57.7121	50
TTCGCTAGACGTACAGGTCA	48	0	43	0	56.3791	50
ATGGAGTCTGTCGAGTAGTG	51	16	54	3	56.5732	50
AAGCGTCACCATAGCCATGA	49	0	47	0	56.5431	50
AGTGGTACGTCCCTGAGTAT	50	0	50	0	56.7012	50
GA						
ATAGAGTGGATAGTTCTGGG	55	9	64	3	52.652 2	45
CATTGGCGGCGCGTAGGCTT	69	0	60	0	69.200 9	65
CTTGTGACCGCTTCTGGGGA	60	16	63	0	60.856 3	60
GAAAAAGGACCAAAAGAGA	58	41	56	0	52.711 1	40
GATGGTGGTTAGAGAAGTGG	58	0	54	0	55.305 6	50
TGTATCTCGTTTTAACATCC	61	16	59	4	48.445 1	35
TTGTAAGCCTACTGCGTGAC	75	0	57	3	56.705 5	50
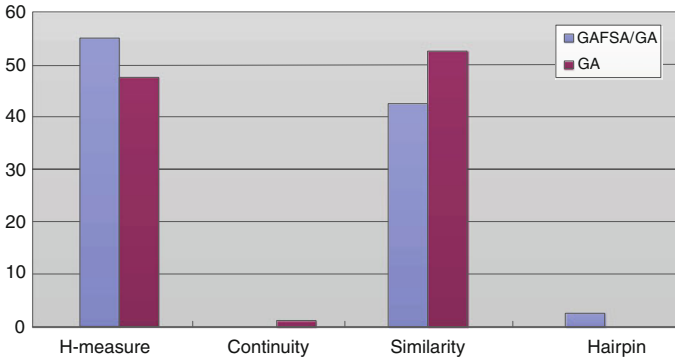
Fig. 1 The compared results of GASFA/GA and GA

5 Conclusion and Prospect

This article from the DNA encode multiple constraint conditions to select the appropriate constraint, the constraint is transformed into a multi-objective optimization problem, put forward GAFSA/GA algorithm on DNA calculation in the coding sequence of the optimization, through the comparison with former, produced better DNA sequence, verified the feasibility and validity of the algorithm. The algorithm the DNA coding quality more stable and reliable, has strong scalability, applicable to a variety of nucleic acid testing of DNA molecular design requirements, can be efficiently produced to meet the coding of constrained maximum DNA sequence set.

References

1. Holland JH (1975) Adaptation in natural and artificial systems [M]. University of Michigan Press, Ann Arbor
2. Deaton R, Murphy RC, Rose JA et al. (1997) A DNA based implementation of an evolutionary search for good encodings for DNA computation[C] In: Proceedings of IEEE conference on evolutionary computation,Indianapolis,IL. Los Alamitos. IEEE Computer Society Press, CA, pp 267–271
3. Wood DH, Chen J (1999) Physical separation of DNA according to royal road fitness[C] In: Proceedings of IEEE conference on evolutionary computation. IEEE Computer Society Press, Washington, pp 1016–1025
4. Fu Y, Zhang D, Xiang X (2011) A combination model to optimize DNA encoding based on global artificial fish swarm algorithm. J Hunan City Univ (Nat Sci) Jun:55–57
5. Cui G, LI X, Zhang XC, YanFeng W (2010) The Optimization of DNA Encodings Based on Modified PSO/GA Algorithm. Chin J Comput, 33:312–313
6. Shin SY, Lee IH, Kim D (2005) Multi objective evolutionary optimization of DNA sequences for reliable DNA computing. IEEE transactions on evolutionary computation, 9(2):143–158

Kernel Optimal Unsupervised Discriminant Projection and Its Application to Face Recognition

Xingzhu Liang, Yu'e Lin and Jingzhao Li

Abstract Kernel Optimal Unsupervised Discriminant Projection (KOUDP) is presented in this paper. The proposed method first maps the input data into a potentially much higher dimensional feature space by virtue of nonlinear kernel trick, and in such a way, nonlinear features is extracted by running UDP on the kernel matrix. The singularity problem of the non-local scatter matrix due to small sample size problem occurred in UDP is avoided. Experimental results on YALE database indicate that the proposed KOUDP method achieves higher recognition rate than the UDP method and other kernel-based learning algorithms.

Keywords Kernel · Unsupervised discriminant projection · Nonlinear features · Kernel matrix

1 Introduction

The feature extraction is a critical issue in face recognition tasks. Typical feature extraction methods are well-known: principal component analysis (PCA), Fisher linear discriminant analysis (FLDA) [1] and locality preserving projection (LPP) [2, 3]. The PCA (known as Eigenface) maps face image into the eigen-space spanned by principal components aiming at minimizing the variance over all face

X. Liang (✉) · Y. Lin · J. Li
School of Computer Science and Engineering, Anhui University of Science and Technology,
Huainan 232001, China
e-mail: lxz9117@126.com

Y. Lin
e-mail: linyu_e@126.com

J. Li
e-mail: jzhli@aust.edu.cn

image samples. Therefore, the PCA method is optimal for data reconstruction, but not for discriminant analysis. The FLDA seeks to find a linear transformation that maximizes the between-class scatter and minimizes the within-class scatter, in order to separate one class from others. Thus, the FLDA can obtain optimal linear discriminant vectors. However, both PCA and LDA fail to discover the underlying structure, if the data live on or close to a sub-manifold of the ambient space. LPP [2–4] is modeled based on the characterization of "locality". The modeling of LPP, however, has no direct connection to classification. The objective function of LPP cannot guarantee to yield a good projection for classification in some cases where the "non-locality" provides dominant information for discrimination. In order to solve the problem, Yang [5] proposed a criterion, which seeked to maximize the ratio of the non-local scatter to the local scatter. This criterion, similar to the classical Fisher criterion, is a Rayleigh quotient in form. Since the proposed method does not use the class-label information of samples in the learning process, this method is called unsupervised discriminant projection (UDP), in contrast with the supervised discriminant projection of LDA. In contrast with LPP, UDP has intuitive relations to classification since it utilizes the information of the "non-locality".

Although UDP seems to be more efficient than other subspace analysis methods for face recognition, it is still a linear technique. Hence it is inadequate to describe the complexity of real face images because of illumination, facial expression and pose variations. Kernel methods [6–9] have been widely used to overcome the limitation of some linear feature extraction and classification and have achieved better results than some linear feature extraction. So we develop a nonlinear unsupervised discriminant projection approach called kernel optimal unsupervised discriminant projection (KOUDP) in this paper. The KOUDP method combines the strengths of both UDP and kernel-based learning techniques to improve the performance of UDP. However, the main difficulty encountered is how to run the eigen decomposition in feature space due to the nonlinear map unknown. To overcome this difficulty, according to the symmetry of kernel scatter matrix with original sample matrix, every column of the kernel matrix is regarded as a corresponding sample, so the discriminant vectors in the kernel matrix can be easily extracted. Then the KOUDP is easily developed in feature space. Finally, experimental results on YALE face database show that the proposed method is effective and feasible.

2 Overview of the UDP

Assume that X is a n dimensional face sample set with N elements belonging to C classes. Denote x_i the ith image and N the number of face sample set. Denote $X = [x_1 \cdots x_i \cdots x_N]$. Defining affinity matrix is as follows:

$$H_{ij} = \exp\left(-\|x_i - x_j\|^2 / t\right) \tag{1}$$

where t represents the neighborhood relation between data sample x_i and x_j. Let S_L be is called the local scatter (covariance) matrix, and S_N is called the non-local scatter (covariance) matrix defined as follows, respectively:

$$S_L = \frac{1}{2NN} \sum_{i=1}^{N} \sum_{j=1}^{N} H_{ij} \left(x_i - x_j\right)\left(x_i - x_j\right)^T \tag{2}$$

$$S_N = \frac{1}{2NN} \sum_{i=1}^{N} \sum_{j=1}^{N} \left(1 - H_{ij}\right)\left(x_i - x_j\right)\left(x_i - x_j\right)^T \tag{3}$$

It is obvious that the sum of S_L and S_N is the total covariance matrix S_T defined as follows:

$$S_T = \frac{1}{2NN} \sum_{i=1}^{N} \sum_{j=1}^{N} \left(x_i - x_j\right)\left(x_i - x_j\right)^T \tag{4}$$

UDP is defined as the following optimization problem

$$w = \arg \max_W \frac{w^T S_N w}{w^T S_L w} \tag{5}$$

If S_L is non-singular, Eq. (6) can be solved by generalized eigenvalue decomposition

$$S_N w = \lambda S_L w \tag{6}$$

The projection axes of UDP can be selected as the generalized eigenvectors $w_1, w_2, \ldots, w_d$ of $S_N w = \lambda S_L w$ corresponding to d largest positive eigenvalues. The concrete procedure of UDP can be fund in [4].

3 KOUDP

Before giving KOUDP, we first perfect the criterion of UDP. As to the criterion of UDP, we first compute $P = (\beta_1 \beta_2 \ldots \beta_m)$, which are non-null space of S_T and corresponding to positive eigenvalues of S_T, and then project face sample set X into the P. But $P^T S_N P$ may be still singular, we first improve the criterion of UDP, which can easily work out the UDP solution no matter whether the $P^T S_N P$ is singular or not. Since $S_T = S_L + S_N$, Eq. (5) is equivalent to minimal criterion as follows:

$$w = \arg \min_{W} \frac{w^T S_L w}{w^T S_T w} \qquad (7)$$

It is proven that $P^T S_T P$ is non-singular, Eq. (7) can be solved by generalized eigenvalue decomposition

$$S_L^P w = \lambda S_T^P w \qquad (8)$$

The projection axes of UDP can be selected as the generalized eigenvectors $w_1, w_2, \ldots, w_d$ of $S_L^P w = \lambda S_T^P w$ corresponding to d smallest eigenvalues.

To extend the linear UDP to the nonlinear case, let X be mapped into a feature space H though a nonlinear mapping Φ. Then we have

$$\Phi : x_i \rightarrow \Phi(x_i) \ \ \Phi(X) = \left[\Phi(x_1)\ldots\Phi(x_i^1)\ldots\Phi(x_N)\right] \qquad (9)$$

In order to derive KOUDP, We first rewrite S_L and S_T in matrix form.

$$S_L = \frac{1}{NN} X L_L X^T \qquad (10)$$

$$S_T = \frac{1}{NN} X L_T X^T \qquad (11)$$

where $L_L = D_L - H_L$, D_L is a diagonal matrix, and its entries are column or row sums of H_L, which is calculated according to Eq. (1). Where $L_T = D_T - H_T$, D_T is a diagonal matrix, and its entries are column or row sums of H_T. According to Eq. (4), It is obvious that entries of H_T are all 1.

Denote the local scatter (covariance) matrix S_L^Φ and the non-local scatter matrix S_T^Φ in the H feature space, respectively. We can obtain

$$S_L^\Phi = \frac{1}{NN} \Phi(X) L_L (\Phi(X))^T \qquad (12)$$

$$S_T^\Phi = \frac{1}{NN} \Phi(X) L_T (\Phi(X))^T \qquad (13)$$

Then we obtain Eq. (7) can be written as

$$w = \arg \min_{W} \frac{w^T S_L^\Phi w}{w^T S_T^\Phi w} \qquad (14)$$

But the nonlinear mapping Φ in Eq. (14) is unknown, non-null space of S_T^Φ can not be achieved directly in the feature space H. This is the main difficulty encountered in developing the algorithm for KOUDP. So we transform Eq. (14) into another form, which only contains the inner product of training samples in the feature space. According to the theory of reproducing kernel [5, 6], any solution w must lie in the span of all the samples in H, so we have

$$w = \Phi(X)a \tag{15}$$

where $a = (a_1, \ldots, a_N)^T \in R^N$, we can get

$$w S_L^\Phi w = a^T K_L a \tag{16}$$

$$w S_T^\Phi w = a^T K_T a \tag{17}$$

where K_L and K_T are the kernel local scatter matrix and the kernel total scatter matrix, respectively, and they are all the $N \times N$ symmetrical matrices. K_L and K_T can be computed by

$$K_L = \frac{1}{NN} K L_L K^T \tag{18}$$

$$K_T = \frac{1}{NN} K L_T K^T \tag{19}$$

where K is a $N \times N$ kernel matrix defined by

$$K = (\Phi(X))^T \Phi(X) \tag{20}$$

where K means to be generated by dot product of all the training samples. Then from Eqs. (16) to (17), Eq. (14) can be written as

$$a = \arg \min_a \frac{a^T K_L a}{a^T K_T a} \tag{21}$$

How to resolve Eq. (21), we present a very simple and efficient method, which only calculates a kernel matrix more than original UDP. From the definitions of K_L, K_T, S_L^Φ and S_T^Φ, we can find that they are very alike, if substituting K for $\Phi(X)$, then S_L^Φ and S_T^Φ become K_L and K_T, respectively. Therefore, each column vector of $\Phi(X)$ is corresponding to each column vector of K. So we take each column vector of K as a sample vector. Equation (21) can be solved by generalized eigenvalue decomposition

$$K_L^V a = \lambda a \tag{22}$$

4 Experimental Results

In order to test performance of our proposed method, we apply it to face recognition. The YALE database is used. The YALE face database is composed of 15 distinct subjects. Each subject has 11 images under different expression and different views.

Table 1 Recognition rates on the Yale face database (%)

Methods	Training samples/class				
	4	5	6	7	8
UDP	70.95	72.67	74.07	74.67	76.56
KOUDP	72.57	73.67	77.87	81.67	83.22
SKLPP	72.43	73.51	76.80	79.17	82.20
KPCA	56.33	58.89	61.60	65.67	69.56

In the experiments, we adopt the Gaussian kernel $k(x, y) = \exp\left(\|x - y\|^2/\sigma\right)$. The nearest distance classifier is adopted due to the simplicity. Table 1 shows the results of the proposed method (KOUDP) and others methods.

All methods are compared on the same training sets and testing sets. The number of training samples per subject k, increases from 4 to 8. In each round, k images are randomly selected from the database for training and the remaining images of the same subject for testing. For each k, 10 tests are performed and these results are averaged. Table 1 shows the average recognition rates. From the experimental results in Table 1, we can find KOUDP outperforms the UDP, KPCA and KLPP method. The one reason is that KOUDP method combines the strengths of both UDP and kernel-based learning techniques to improve the performance of UDP. The other is that the criterion used in the KOUDP is optimal compared with that used in the KLPP and KPCA.

5 Conclusions

In this paper, we present the algorithm called KOUDP by combining the UDP with kernel method. Considering that $P^T S_N P$ may be singular, we first perfect the criterion of UDP and then proposed KOUDP. According to the definitions of K_L, K_T, S_L^Φ and S_T^Φ, each column vector of K is considered a corresponding sample. Then, the KOUDP is easily developed in kernel feature space. Finally, experimental results YALE face databases show that the proposed method is effective and feasible.

References

1. Belhumeur PN, Hespanha JP, Kriengman DJ (1997) Eigenfaces vs. fisherfaces: recognition using class specific linear projection. IEEE Trans Pattern Anal Mach Intell 19(7):711–720
2. He XF, Yan SC, Hu Y et al (2005) Face recognition using Laplacianfaces. IEEE Trans Pattern Anal Mach Intell 27(3):328–340
3. He XF, Niyogi P (2004) Locality preserving projections. Proc Adv Neural Inf Process Sys 16:153–160 (MIT Press, Cambridge)
4. Zhang T, Huang K, Li X, Yang J, Tao D (2010) Discriminative orthogonal neighborhood-preserving projections for classification. IEEE Trans Syst Man Cybern Part B Cybern 40(1):253–263
5. Yang J, David Z, Yang JY, Ben N (2007) Globally maximizing, locally minimizing: unsupervised discriminant projection with applications to face and palm biometrics. IEEE Trans Pattern Anal Mach Intell 29(4):650–664
6. Scholkopf B, Smola A et al (1998) Nonlinear component analysis as a kernel eigenvalue problem. Neural Comput 10(5):1299–1319
7. Cheng J, Liu QS, Lu HQ et al (2005) Supervised kernel locality preserving projections for face recognition. Neurocomputing 67:443–449
8. Li JB, Pan JS, Chu SC (2008) Kernel class-wise locality preserving projection. Inf Sci 178(7):1825–1835
9. Li JB, Pan JS, Chu SC (2011) Kernel self-optimized locality preserving discriminant analysis for feature extraction and recognition. Neurocomputing 74(17):3019–3027

The Binary Anti-Collision Algorithm Based on Labeled Packet

Jianbin Xue, Lijing Qin and Wenhua Wang

Abstract An improved algorithm based on label packet (PABI) is put forward, which aimed at much times search and the large amount of communication data of current binary anti-collision algorithm. The new algorithm improve decoding accuracy by grouping labels within the field that reader identify by setting a counter on the label, extract conflicts bit then make up conflict block in the recognition process, and introduce the concept of matrix to express the possible decode of conflict block, thereby. The simulation results show that the algorithm have a significant advantage in the number of searches and the amount of communication data in the case of a large number of tags.

1 Introduction

Radio Frequency Identification (RFID) is a noncontact automatic identification technology [1]. The existing anti-collision algorithm can be divided into two categories: one is based on the ALOHA (ALOHA algorithm [2], slot ALOHA algorithm [3], the dynamic slot ALOHA algorithm [4]), which is a uncertain probabilistic algorithms. The other is tree-based anti-collision algorithm (basic

---

J. Xue (✉) · L. Qin · W. Wang
School of Computer and Comunication, Lanzhou University of Technology,
730050 Lanzhou, Gansu, People's Republic of China
e-mail: Xuedy369@126.com

L. Qin
e-mail: qlj19890108@163.com

W. Wang
e-mail: wwhlzlg@126.com

binary search algorithm [5], dynamic binary search algorithm [6]), this algorithm can solve the problem of "label hunger" and improve recognition accuracy and the speed of recognition, so many scholars have studied tree-based anti-collision algorithm. The above studies improve the binary anti-collision algorithm, but there is still redundancy in the amount of communication between the reader and the labels, the number of interactions is still large.

2 Improved Binary Anti-collision Algorithm

Compared with the dynamic binary search algorithm, the improvements of PABI algorithm are mainly in the following two aspects:

First improvement: Set a counter T called the label packet counter that only value 0 and 1 within the label. Before the reader decode all the labels within the recognition field, a pretreatment is conducted to response label, that is, computing the label ID as phase "and", the results are stored in their own counter T, and then all labels are divided into two groups (T = 1 and T = 0), so the value of T indicates whether the number of 1 in the label ID is odd or even (when T = 1, is even; when T = 0, is odd number), and the subsequent identification process in accordance with the principle that label group that T = 1 is prior, thus we can avoid part of conflict and reduce the amount of data that reader once handle.

Second improvement: Using matrix to represent the possible values of the conflict block, this matrix is called the conflict block matrix P. The correctly receiving portion of reader according to that the results of bitwise "AND" refer as R, the same values of R can only be 0 and 1. After pretreatment, when the answer signal of reader decoding conflict, the number of conflict bits is extracted to compose the conflicting block, and count the number of conflict bits through the conflicting block bit-digit counter Q, so the value of conflicting block matrix is correlate to the value of Q.

2.1 Process of the PABI Algorithm

Let the ID of labels to eight. And assume that label packet group that counter T = 1 has higher priority than the T = 0 group.

(1) When the reader sends request command, all the labels that ID code value is less than or equal to UID should response to this command, and send its ID code.
(2) ID of the response label within reader field is divided into two groups according phase "and", and phase "and" results is stored in their own packet counter T, they can be divided into two groups of T = 1 and T = 0.
(3) Decoding to the response signal that T = 1, determine whether there is a conflict according to the decoding results. If there is no conflict, reader sends the

Select command and Read-Data to read data, finally sends Unselect command to let the label enter to choose state. If there is a conflict, then detect the conflict bit and extract to compose conflict module (the length of the module is equal to the conflict bit number), and calculate the result of phase "and" correct receiver part, record as R.

(4) If the conflict block bit count Q = 1(the number of conflict bit is only one), then conduct the conflict bock matrix, labels can be directly identified. Otherwise, go to step (5).

(5) If Q = 2 (two conflicts), compare the values of T and R, if equal, indicating that the parity of 1 of correctly receiving portion is same to conflict block, so conflict block matrix; if not equal indicating that there is one bit 1 in the conflict block, so conflict block matrix. If Q ≠ 2, then go to step (6).

(6) If Q = 3 (three conflicts), compare the values of T and R, if equal, the conflict block matrix; if not equal, the conflict block matrix. If Q ≠ 3, then go to step (7).

……

Until all the labels of T = 1 group are identified. Then go to step (n).

(n). Decoding to the response signal of T = 0 group. The following execution order is same to the T = 1 group.

……

Until all the labels of T = 0 group are identified. All the identification processes are finished.

3 Performance Analysis

3.1 The Number of Searches of the Reader

If there are N number labels in the recognition range of the reader, assume that the number of conflict is M, the number of searches of the general binary algorithm is 2N−1. The search number comparison of this algorithm and general binary search algorithm is shown in Table 1. (N ≥ 2, assume that there are three conflicts at most). When N = 1, there is only one label in the identification field, the number of search time is 1.

Table 1 Number of the reader searches of the general binary and PABI algorithm

/	N	/	2	3	4	5	6	7	8	9	10	11	12
General	/	2 N−1	3	5	7	9	11	13	15	17	19	21	23
PABI	M = 1	N + 2	4	5	6	7	8	9	10	11	12	13	14
	M = 2	N + 2	4	5	6	7	8	9	10	11	12	13	14
	M = 3	N + 4	6	7	8	9	10	11	12	13	14	15	16

As can be seen from the comparison, when the number of labels increasing, the search times of the PABI algorithm is less than the general binary algorithm.

3.2 The Amount of Data

Let the number of labels is N, and the length of the label ID is x, and the distribution of the label is uniform.

(1) The amount of data sent by the reader.

General algorithm: the reader identifies the N label need 2N-1 time slots, i.e., the amount of data sent by the reader:

$$L_R = (2N - 1)x \tag{1}$$

The PABI algorithms: when the collision median M = 1 or 2, identify the N labels need N + 2 slots, the amount of data sent by the reader:

$$L'_{R1} = (N + 2)x = L'_{R2} \tag{2}$$

When M = 3, the amount of data sent by the reader as follows:

$$L'_{R3} = (N + 4)x \tag{3}$$

(2) The amount of data sent by label.

General algorithm: Identification of the N labels that experienced maximum collision need N−1 slot, then the total data amount of query command that the label response to the reader for the first time is NX, when there is a collision, the back strategy is used to identify each branch, then the amount of data by labels varies with the search depth, the total amount of data sent by the label, Where C is a back number:

$$L_T = Nx + Cx \tag{4}$$

The PABI algorithms: regardless of the median of the collision, we can find out the accurate decoding in the conflict block matrix through the judgment and the comparison and then send out, because the collision median is far less than the median of label ID, so post-conflict only return reader collision bit decoding, that is, several pulses. The amount of data sent by the label is as follows, Where M(n) is the number of pulses sent each time:

$$L'_T = Nx + M(n) \tag{5}$$

4 Simulation and Analysis

Assume that the channel is ideal. Let the label length is 64 bit, the label amount varies from 0 to 100, Fig. 1 is the comparison of the reader search times of PABI algorithm and BLBO algorithm [7] when M = 1, 2, 3.

As Fig. 1 show, with the conflict median increasing, the reader search amount are also increasing, but with the growing of the number of labels, the reader search time is obviously less the BLBO algorithm, compare to the BLBO algorithm reader search time, the influence that the increase of the PABI algorithm conflict amount to the search time is very little.

As shown in the Fig. 2, when the conflict median increase in the PABI algorithm, the amount of data sent by readers increase slightly, but it is significantly lower than the data amount of general binary algorithm. And with the increase in
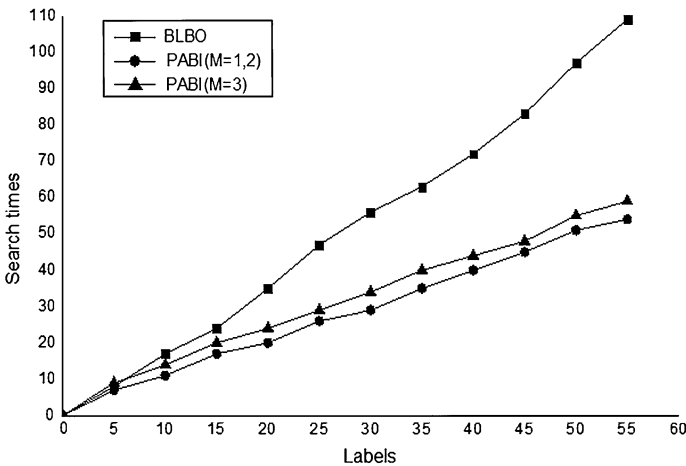
Fig. 1 The comparison of the reader search times of PABI algorithm and BLBO algorithm
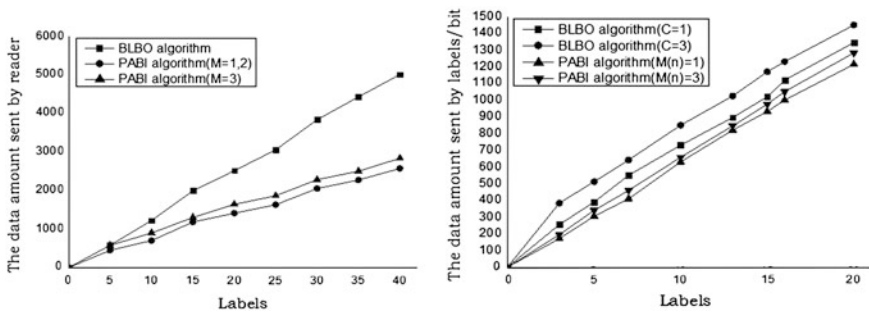
Fig. 2 The comparison of the data amount sent by reader and labels of PABI algorithm and BLBO algorithm

the number of backward, the amount of data sent by general binary algorithm labels increases, as the conflict bit number in PABI algorithm increases, the amount of data the label transmitted increased slightly, i.e., in PABI algorithm, the power consumption of the label is less than binary algorithm.

5 Conclusion

A PABI algorithm is proposed in this paper, which packets the labels within identification range by setting label packet counter, and in conflict recognition process, the reference of matrix concept to represent the possible decoding of the conflict block composed of the conflict bit extracted from the labels, make its search accuracy greatly improved. The result of the analysis and simulation of the performance of the algorithm show that PABI algorithm based on label packet are significantly better than general binary algorithm in the reader search times and communication data amount.

References

1. Finkenzeller K (2003) RFID—handbook fundamentals and applications in contactless smart cards and identification, 2nd edn. Wiley, New York
2. Tanenbaum AS (2003) Computer networks, 4th edn. Prentice Hall, New Jersey
3. Bonuccelli MA, Lonetti F, Martelli F (2006) Tree slotted aloha: a new protocol for tag identification in RFID networks. In: Proceedings of IEEE international symposium on a world of wireless, mobile and multimedia networks, pp 603–608
4. Cha JR, Kim JH (2006) Dynamic framed slotted ALOHA algorithms using fast tag estimation method for RFID system. Consumer communications and networking conference, CCNC 2006, 3rd IEEE, pp 768–772
5. Jiang C, Huang L (2011) Study on tag anti-collision algorithms of RFID based on binary search. Comput Dig Eng 39(4):30–31
6. Yu SS, Zhan YJ, Wang ZP et al (2005) Anti-collision algorithm based on jumping and dynamic searching and its analysis. Comput Eng 31:19–20
7. Wang X, Qian ZH, Hu ZC (2010) Research on RFID anti-collision algorithms based on binary tree. J Commun 31(6):49–57

Research and Implementation on Event-Based Method for Automatic Summarization

Tao Liao, Zongtian Liu and Xianchuan Wang

Abstract By studying the technology of automatic summarization, this paper considers event as a basic semantic unit for narrative texts, and presents a new Event-based method for automatic summarization. This method utilizes events and the relationship between events to build Event-Network text representation model, which can retain the structure information and semantic information of the text to a greater extent. The experimental results show that the Event-based automatic summarization method has better performance.

1 Introduction

The world is material, the material world is dynamic. Event originated from cognitive science, often appears in the texts of philosophy, cognitive science, linguistics and Artificial Intelligence. Cognitive scientists believe that humans

T. Liao (✉) · Z. Liu · X. Wang
School of Computer Engineering and Science, Shanghai University, 200072 Shanghai, China
e-mail: tliao@aust.edu.cn

Z. Liu
e-mail: ztliu@shu.edu.cn

X. Wang
e-mail: xch_wang@shu.edu.cn

T. Liao
School of Computer Science and Engineering, Anhui University of Science and Technology, 232001 Huainan, China

memorize and understand the real world through "event". In recent years, the concept of "event" has been widely used in the computational linguistics as well as artificial intelligence, information retrieval, information extraction, automatic summarization and other natural language processing applications.

As the basic unit of human knowledge, the event reflects the movement, behavior and changes in the real world. The event associates with some of the concepts such as actors, time, place and so on, and is a greater semantic unit than the concept. A large number of narrative class texts contain various events, for example, novels, biographies, news reports and other narrative class texts. We consider the event as a basic semantic unit for these narrative texts, and establish a new Event-based text representation method (Event-Network) to understand the semantics of sentences and articles. By studying the technology of automatic summarization, this paper presents a new Event-based method for automatic summarization according to the characteristics of narrative texts.

2 Automatic Summarization

Research on automatic summarization started in the 1950s. As early as 1952, H. P. Luhn began to study the method of automatic summarization and published the first article on the methods of automatic summarization [1]. Automatic summarization can be roughly classified into three categories.

1. Corpus Based Methods. By exploiting technologies of machine learning, Corpus-Based Methods [2] become possible to learn rules from a corpus of documents and the corresponding summaries. The process is decomposed into two phases: the training phase and the test phase. In the training phase, the system extracts particular features from the training corpus and generates rules by a learning algorithm. In the test phase, the system applies rules learned from the training phase to the test corpus to generate the corresponding summaries.
2. Statistical model Based Methods. Statistics-Based Methods use some of the statistical features or semantic features to extract important sentences, thus generate a summary. Early methods judge the importance of sentence mainly on the basis of term frequency, title, location, sentence structure and clue words. Later, some Domain Ontology-Based Methods appeared [3]. These methods analyze the document according to semantic information and internal relations of domain ontology, filter out irrelevant information, and can extract more accurate summary.
3. Graph Structure Based Methods. Graph Structure Based Methods [4] extract the characteristic words, sentences, paragraphs or other text features as nodes in the graph. Each node stands for a feature and a link is created between two nodes if the two corresponding features have strong relevance. Graph Structure Based Methods use graph to represent text or text set, and utilize graph theory to judge

the importance of text features, and then calculate the weights of sentences or paragraphs.

Corpus Based Methods apply rules learned from the training phase to the test corpus to generate the corresponding summaries, which excessively depend on the training corpus and are difficult to ensure the accuracy of the summary; Statistical model Based Methods lack the analysis of the topic and content of the text, and need to take a long time to build their own domain ontology because of the lack of Chinese domain ontology; Compared with the previous two methods, Graph Structure Based Methods can reflect the structural relationship and semantic links between text features, and greatly improve the quality of the summary.

3 Event and Event Similarity

Different applications define event in very different ways. In the field of information retrieval, event is considered as a refinement of the search for the theme [5]. In automatic summarization, Filatova et al. [6] defined atomic events as combination of major constituent parts of the actions and the event term verbs or action- nouns which are the links of constituent parts. An event is expressed as a triple {name entity, event term, name entity}. Different from Filatova's triple, Zhou Wen [7] proposed an event multi-tuple model. Zhou's model didn't limit the number of constituent parts to 2, but extracting constituent parts according to the existence in the sentence. From the above, although the definitions of event are not unified in different applications, most of them emphasize two kinds of event attributes, action (verb or action-noun) and action characteristics (participant, location, time, etc.).Here, we give the following definitions of event.

Definition 1 (*Event*) We define *event* as a thing happens in a certain time and environment, which some actors take part in and show some action features. Event e can be defined as a 6-tuple formally:

$$e = (A, O, T, V, P, L)$$

We call elements in 6-tuple event factors. A means an action set happen in an event. O means objects take part in the event. T means the period of time that event lasting. V means environment of event. P means assertions on the procedure of actions execution in an event. L means language expressions.

Event is composed of event elements, and we can obtain four event elements directly from the text, including the action A, the object O (subject and object), the time T, and the environment V. Thus we can make use of these four elements of the event to calculate the similarity between two events.

Definition 2 (*Event similarity*) *Event Similarity* of the two events e_m and e_n is denoted by $Sim(e_m, e_n)$, defined as follows:

$$Sim(e_m, e_n) = \sum_{i=1}^{4} w_i s(x_{im}, x_{in}) \; x_i \in \{A, O, T, V\}$$

x_{im}, x_{in} represent event elements of e_m and e_n respectively, such as action, object, time, environment of event. $s(x_{im}, x_{in})$ represents the Similarity of the two events e_m and e_n. The values are as follows: (1) If two elements are synonyms, such as the environmental element of e_m is "上海", and the environmental element of e_n is "沪", $s(x_{im}, x_{in}) = 1$; (2) If two elements have the relationship of inclusion or hyponymy, such as the time element of e_m is the evening of July 23, and the time element of e_n is at 20:34 on July 23, $s(x_{im}, x_{in}) = 0.5$; (3) In other cases, $s(x_{im}, x_{in}) = 0$. w_i is the weighting factor of event elements, subject to $\sum w_i = 1$. Setting the threshold value α, we can think that e_m and e_n are similar when $Sim(e_m, e_n) \geq \alpha$.

Considering the description ability of various event elements, we set $w_1 = 0.5$, $w_2 = 0.3$ and $w_{3,4} = 0.1$ in this paper. By experimental observation, we think that e_m and e_n are similar when the value of α is 0.7.

4 Event-Based Automatic Summarization

This paper considers the event as a basic semantic unit for narrative texts, and presents a new Event-based method for automatic summarization. This method combines the characteristics of the graph structure to build the Event-Network text representation model, and extracts summary on this basis.

4.1 Annotation of Events

This paper firstly downloads 200 emergency news texts from the Internet as an experimental corpus, and this corpus includes five event classes: earthquake, traffic accident, terrorist attack, food poisoning and fire. Then we have developed an event annotation tools, and use XML language to label each event and its event elements of the text. Our corpus is called CEC (Chinese Emergency Corpus). The effect diagram of labeled events is shown in Fig. 1.

4.2 The Definition of the Event-Network

The Event-Network is an undirected graph, which is triples of the form $G = (N, E, W)$, where

N The nodes set $N = \{n_1, n_2, \ldots, n_i, n_j, \ldots, n_k\}$, each node n_i represents an event feature, and k is the number of nodes of the graph structure

<Event eid="e1"><Time type="absTime">2009年4月14日
</Time> <Location>杭州-上海高速公路</Location>上发生一起
<Denoter type="emergency">交通事故</Denoter> </Event>，
<Event eid="e2">导致<Participant>4 人</Participant> <Denoter
type="stateChange">死亡</Denoter></Event>，<Event
eid="e3"><Participant>2人</Participant><Denoter
type="stateChange">受伤</Denoter></Event>。

Fig. 1 The effect diagram of labeled events

E The set of undirected edges $E = \{\ldots, e_{ij}, \ldots\}$, each undirected edge e_{ij} (i, $j = 1, 2, \ldots, k$, and $i \neq j$) represents the neighboring relationship in the paragraph or event similarity of the two corresponding event features (n_i and n_j)

W The weights set of nodes $W = \{w_1, w_2, \ldots, w_i, w_j, \ldots, w_k\}$, w_i is the weight of the node n_i, and it is calculated as follows:

$$w_i = d_i k / \sum_{m=1}^{k} d_m \ (i = 1, 2, \ldots, k)$$

d_i represents the degree of the node n_i, k is the total number of nodes in the Event-Network.

4.3 Construction of the Event-Network

This paper firstly uses the self-development annotation tools to manually label events of the text d, and extracts events and their event elements from the labeled text d to get the event features set $EV_d = \{e_1, e_2, \ldots, e_i, e_j, \ldots, e_k\}$, and then constructs the Event-Network on this basis.

Construction of the Event-Network includes the following steps:

(1) Initializing Event-Network, the nodes set $N_d = \{\}$, the undirected edges set $E_d = \{\}$, and the weights set of the nodes $W_d = \{\}$;
(2) Mapping event features of the event features set $EV_d = \{e_1, e_2, \ldots, e_i, e_j, \ldots, e_k\}$ to nodes of the Event-Network graph structure, and getting the nodes set $N_d = \{n_1, n_2, \ldots, n_i, n_j, \ldots, n_k\}$;
(3) Selecting two adjacent nodes n_i and n_j from N_d, and adding one undirected edge e_{ij} if their corresponding events are in same paragraph of the text d;
(4) Taking any two nodes n_i and n_j from N_d, and adding one undirected edge e_{ij} between n_i and n_j if their corresponding events are similar; Getting the undirected edges set $E_d = \{\ldots, e_{ij}, \ldots\}$ from (3) and (4).

(5) Counting the degree of each node in the Event-Network, and calculating the weight of each node according to the node weight formula.

Follow the steps above, we can get the nodes set $N_d = \{n_1, n_2, …, n_i, n_j, …, n_k\}$, the set of undirected edges $E_d = \{…, e_{ij}, …\}$, and the weights set of nodes $W_d = \{w_1, w_2, …, w_i, w_j, …, w_k\}$, thus construct the Event-Network of the text d.

4.4 An Example

This paper uses an example to specify the construction process of the Event-Network. A Chinese news text d titled"18名路人见死不救 拾荒阿姨救起遭碾压女童 (18 passers do nothing about the traffic accident girl, until a scavenger aunt saves her)", downloaded from the http://news.qq.com/a/20111017/000066.htm.

The text d contains 7 paragraphs and 21 sentences. We use the annotation tools to label the events of the text d, and generate the undirected edges by utilizing the Event-Network construction method proposed in this paper, thus obtain 57 nodes and 88 undirected edges. We apply the visualization software *NetDraw* to visualize the text d, and each node is labeled by the corresponding event and its action element. The Event-Network representation of the text d is shown in Fig. 2.

Obviously, the Event-Network can not only contain the information of event features, and contain the text structure information and semantic similarity information between event features. Therefore, the Event-Network can better represent the knowledge of the text.
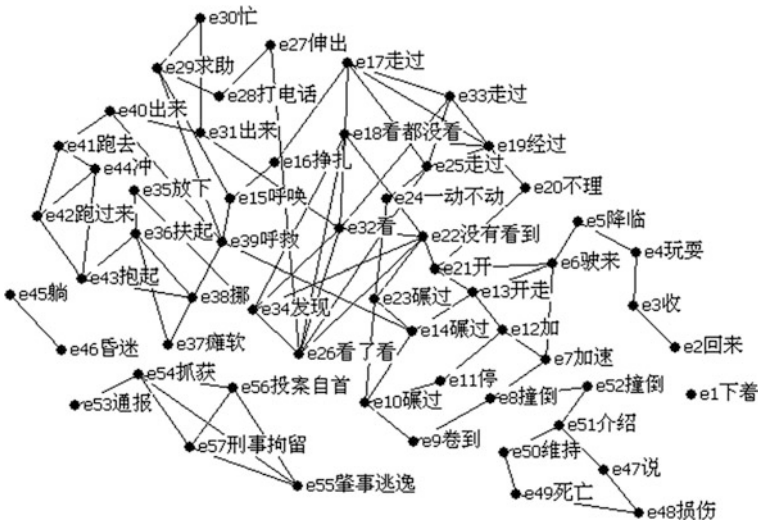
Fig. 2 The event-network representation of the text d

4.5 Automatic Summarization Generation

Automatic summarization regards the text as the linear sequence of sentences, and sentence as a linear sequence of events, specifically includes three steps:

1. Using weights of event features to calculate the weight of each sentence.
2. All the sentences of the original are in descending order according to their weight, and some sentences are identified as summary sentences because of their higher weight.
3. Outputting all summary sentences according to their appearance order in the original.

5 Experiments and Discussion

In this section, we report our preliminary experimental results.

On the evaluation criteria, we used precision P, recall R and F to judge the quality of the summary. We Assume that x is the manual summary and y is the machine generated summary, $P = (x \cap y)/y$, $R = (x \cap y)/x$, $F = 2rp/(r + p)$

In the experiment, this paper randomly selected 10 texts from each event class of the CEC, and totally got 50 texts from five event classes. We respectively used *Microsoft Word* auto-summary method and our method to generate summaries (compression ratio is about 30 %).The experimental result is shown in Table 1.

Table 1 shows that the effect of our method is better than *Microsoft Word* auto-summary method. In addition, this paper compared the experimental results with some other research results at home and abroad, and our comparison results are listed in Table 2.

As can be seen from Table 2, our method has some limitations (only applies to some narrative class texts that contain various events) compared to domestic and foreign automatic summarization methods, but also achieve good results.

Table 1 Experimental comparison of our method and *MS Word*

Class	MS word			Our method		
	P	R	F	P	R	F
Earthquake	0.48	0.42	0.45	0.62	0.56	0.59
Traffic accident	0.52	0.47	0.49	0.66	0.55	0.60
Terrorist attack	0.54	0.46	0.50	0.68	0.60	0.64
Food poisoning	0.45	0.43	0.44	0.67	0.57	0.62
Fire	0.55	0.51	0.53	0.63	0.55	0.59
Average	0.51	0.46	0.48	0.65	0.57	0.61

Table 2 The results of our method compared with other studies

Author	Experimental data	Experimental results		
		P	R	F
Wang [8]	2003 national 863 automatic summarization system evaluation test set, a total of 10 documents	0.60	0.59	0.59
Ge B [9]	27 News articles, 20 literature articles, and 20 scientific papers, a total of 67.	0.57	0.54	0.56
Jaruskulchai [10]	10 articles selected from the Ziff-Davis corpus	0.60	0.44	0.50
Our method	Randomly selected 10 texts from each event class of the CEC, a total of 50 texts	0.65	0.57	0.61

6 Conclusion

In this paper, we consider the event as a basic semantic unit according to narrative texts, and present a new Event-based method for automatic summarization by studying the technology of automatic summarization. This method utilizes events and the relationship between events to build the Event-Network text representation model, and extracts summary on this basis. The Event-Network is more suitable for the objective law of knowledge representation, and it can better represent the semantic information of the text. The experimental results show that Event-Based automatic summarization method has better performance.

References

1. Luhn HP (1958) The automatic creation of literature abstracts. IBM J Res Dev 2(2):159–165
2. Hovy E, Lin CY (1999) Automated text summarization in SUMMARIST. In: Mani I, Maybury M (eds) Advances in automated text summarization. MIT Press, Cambridge, pp 81–94
3. Wang M, Xu C, Li CG, He TT (2011) Method of multi-document automatic summarization based on sub-topic area partition. Comput Eng 37(12):158–160, 163
4. Mihalcea R, Tarau P (2004) TextRank: bringing order into texts. In: Proceedings of the conference on empirical methods in natural language processing. Barcelona, Spain, pp 404–411
5. Yu M, Luo WH, Xu HB, Bai S (2006) Research on hierarchical topic detection in topic detection and tracking. J Comput Res Dev 43(3):489–495
6. Filatova E, Hatzivassiloglou V (2003) Domain independent detection, extraction, and labeling of atomic events. In: Proceedings of RANLP, Borovetz, Bulgaria, pp 145–152
7. Zhou W, Liu Z T (2007) Event-based knowledge acquisition for ontology learning. In: Proceedings of the 6th IEEE international conference on cognitive information, Lake Tahoe, California, USA, pp 498–501

8. Wang ZQ, Wang YC, Liu CH (2007) A sentence weighting method for automatic text summarization based on MRP. J Shanghai Jiaotong Univ 41(8):1297–1300
9. Ge B, Li FF, Li F, Xiao WD (2011) Subject sentence extraction based on undirected graph construction. Comput Sci 38(5):181–185
10. Jaruskulchai C, Kruengkrai C (2003) Generic text summarization using local and global properties of sentences. In: Proceedings of the IEEE/WIC international conference on web intelligence. IEEE Press, Piscataway, pp 201–206

Closed Circle DNA Algorithm of Maximum Weighted Independent Set Problem

Qingyan Li, Zhixiang Yin and Min Chen

Abstract Closed circle DNA algorithm of maximum weighted independent set problem is proposed upon closed circle DNA computing model and its biochemistry experiment. In the algorithm, first we get all independent sets though an appropriate encoding and delete experiments, and then we find the maximum weighted independent set using electrophoresis experiment and detect experiment. Only using delete experiment, the algorithm is simple and credible.

Keywords DNA computing · Closed circle DNA computing model · Maximum weighted independent set problem

1 Introduction

In 1994, Adleman [1] used annealing and connection of linear DNA molecules and worked out Hamiltonian Path Problem (HPP) of a directed graph. He first proposed DNA computing method, which had opened a new field in solving the NP complete problems. Later, the researchers have devoted substantial effort to this new field and have made some progress.

DNA computing model is the main content of DNA computing research, there are many DNA computing models proposed, closed circle DNA computing model

Q. Li · Z. Yin (✉) · M. Chen
Anhui University of Science and Technology, Huainan 232001, China
e-mail: zxyin66@163.com

Q. Li
e-mail: comeon@163.com

M. Chen
e-mail: chenmin0925@163.com

is a new model of DNA computing. This model has had a successful application in solving many problems [2–6].

There are many algorithms to solve non-weighted maximum independent set problem [7–10]. In this paper, upon closed circle DNA computing model, an algorithm of the maximum weighted independent set problem [11] is proposed. We first introduce closed circle DNA computing model applied in the algorithm, and then give the closed circle DNA algorithm and its realization, last show an instance of the algorithm.

2 Closed Circle DNA Computing Model

Closed circle DNA is connected by recognition sequences of several restriction enzymes which connect the linear DNA end-to-end in the recognition site. Given $p_i(i = 0, 1, 2, \ldots, n)$ is the recognition sequence of $n + 1$ different restriction enzymes, each p_i is indicated by $s_i t_i$, and the recognition site is between s_i and t_i. The structure of recognition sequence is shown in Fig. 1. Connect $n + 1$ recognition sequences end-to-end and form the shortest closed circle DNA molecule l^0, which is shown in Fig. 2.

The biochemistry experiments used in this model are as follows:

1. Insert experiment: Insert foreign DNA in the recognition sequence of closed circle DNA molecule. The structure of foreign DNA molecule must be $a_i = t_i b_i s_i$. Add restriction endonuclease i into closed circle DNA molecule l^0, open closed circle DNA molecule between s_i and t_i, and generate linear DNA molecule. Delete restriction endonuclease after doing high performance liquid chromatography (HPLC), by the action of T_4 DNA ligase, linear DNA molecule return to closed circle DNA.

We recognize linear DNA molecule inserted in the different position of closed circle DNA according to the order of the recognition sequence. If linear DNA $a_1, a_2, \ldots, a_n$ is successively inserted in recognition sequence $p_1, p_2, \ldots, p_n$ of closed circle DNA, the closed circle DNA is indicated by $(a_1 a_2 \ldots a_n)$. If no linear DNA is inserted in recognition sequence p_i, where is indicated by 0.

2. Delete experiment: experiment: By the action of restriction endonucleases, some linear DNA molecules are cut from all the closed circle DNA molecules. By the action of restriction endonucleases i, closed circle DNA molecules are changed into a_i and linear DNA molecules which don't contain a_i.

Fig. 1 The structure of p_i
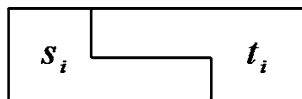
Fig. 2 The structure of
closed circle DNA l^0
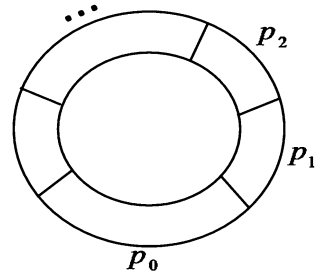
3. Electrophoresis experiment: Use gel electrophoresis, separate DNA molecules according to their length and delete a_i and restriction endonucleases. Finally by the action of T_4 DNA ligase, form new closed circle DNA mixture, and each closed circle DNA molecule doesn't contain a_i.
4. PCR amplification: By the action of polymerase Taq DNA, select 3 temperatures for thermal cycling to amplify DNA molecules. The operation will open circle DNA into linear DNA, select p_0 for the open site to protect recognition sequences $p_1, p_2, \ldots, p_n$, so set up a recognition site in p_0.
5. Detect experiment: Read the DNA sequences expressed solutions from the final reaction product. First, open closed circle DNA into linear DNA, then fix them on the chip to get DNA chip. Last do mark experiment. Mix the marker probes with DNA chips sufficiently to hybridize, wash probes molecules not hybrid, then scan DNA chips with laser focusing microscope to detect the hybridization result and record fluorescence light dot of DNA chips.

3 Resolving the Maximum Weighted Independent Set Problem with Closed Circle DNA Computing Model

3.1 The Maximum Weighted Independent Set Problem

Given a simple undirected weighted graph $G = (V, E, W)$, $V = \{v_1, v_2, \ldots, v_n\}$ is the set of vertices, $E = \{e_1, e_2, \ldots, e_m\}$ is the set of edges and $W = \{w_1, w_2, \ldots, w_n\}$ is the set of weights. An independent set of G is a subset $V' \subseteq V$ such that for any vertex $v \in V'$, between any vertices are not adjacent. If V' is an independent set, and G has no other independent set K which satisfies $|K| < |V'|$, then V' is called a maximum independent set of G. If V' is an independent set, and the sum of all vertex weights is the maximum, then V' is called a maximum weighted independent set of G.

3.2 Closed Circle DNA Algorithm

Any weighted independent set problem of positive weighted graph $G = (V, E, W)$ can be converted to that $G' = (V, E, W')$ with positive-integer weighted as small as possible [3]. Here we only discuss positive-integer weighted independent set problem. In this paper, we propose the closed circle DNA algorithm of the maximum weighted independent set problem, to do that:

Step1. Encode each vertex and produce three groups DNA encodings;
Step2. Construct the initial closed circle DNA mixtures;
Step3. Produce all independent sets of the graph G though delete experiment;
Step4. Produce the maximum weighted independent set using electrophoresis experiment;
Step5. Detect and output results.

3.3 The Realization of Closed Circle DNA Algorithm

3.3.1 The Biochemistry Process of Step 1

1. Encode each vertex of the graph G and produce three groups DNA encodings:

The first group (vertex): $3' - p_0 p_1 - 5' = p_0 s_1 t_1$, $3' - p_i - 5' = s_i t_i$ $(i \neq 1)$, $|p_i| = 6 byte$;
The second group (weight): $3' - a_i - 5' = t_i b_i s_i$, $|a_i| = 6 + |b_i| = 12 \times w_i$;
The third group (marker probe): $5' - a'_t - 3' = t_i b'_t s_i$.

Encodings of the first group represent vertex v_i and recognition sequence p_i, and p_0 is the opening circle recognition sequence. Encodings of the second group represent the weight of vertex v_i. Encodings of the third group represent marker probes. a'_i is the complementary sequence of a_i, and is made with fluorescent;

2. Synthesize three groups DNA coding molecules and use PCR amplification to get enough DNA mixtures.

3.3.2 The Biochemistry Process of Step 2

1. Synthesize deck group: $5' - t_n p_0 - 3'$, $5' - p_1 s_2 - 3'$, $5' - t_i s_i - 3'$, $(i = 2, 3, \ldots, n - 1)$;
2. For $3'$ end of $3' - p_0 p_1 - 5'$, $5' - t_n p_0 - 3'$ and $5'$ end of $3' - p_n - 5'$, $5' - p_1 s_2 - 3'$ do activation disposal to lose chemistry;

3. Take equal molar mass the deck group and the first encodings group and mix fully, do the hybridization experiment by the action of T_4 DNA ligase;
4. Save DNA whose chain length is 6(n + 1), delete T_4 DNA ligase and DNA which do not react, remove the modified protection base at both ends, wash impurities with high performance liquid chromatographic method. Do the hybridization experiment by the action of T_4 DNA ligase, all linear DNA turn into closed circled DNA, and wash impurities; Let $i = 1$;
5. By the action of restriction endonuclease corresponding to p_i do insert experiment to insert $3' - a_i - 5'$;
6. If $i = n$, and return to 3, else let $i = i + 1$, repeat 5.

There is a linear DNA fragment in each closed circle DNA in obtained mixture.

3.3.3 The Biochemistry Process of Step 3

1. Let $i = 1$;
2. For edge $e_i = v_j v_k$ ($j, k = 1, 2, ...,n$), take out two kinds of restriction endonuclease corresponding to two vertices v_j and v_k of e_i, divide the mixture into two equal parts, for each part do delete experiment respectively by the action of restriction endonuclease;
3. Mix two parts after doing delete experiments into together;
4. If $i = m$, and obtain all independent sets of the graph G, return to step 4; else let $i = i + 1$, and return to 2.

3.3.4 The Biochemistry Process of Step 4

1. By the action of restriction endonuclease corresponding to the recognition sequence p_0, all closed circle DNA of the mixture are opening into linear DNA;
2. Under ultraviolet irradiation, cut the zone of gel where the DNA molecular is the longest. For the zone of gel, do gel electrophoresis experiments to free DNA molecule into the buffer solution. So obtain the longest linear DNA mixture, which represents the set of all maximum weighted independent sets.

3.3.5 The Biochemistry Process of Step 5

1. Raise temperature of the linear DNA mixture to 95 °C to react degeneration, and then fall rapidly to 68 °C so as to obtain single-strand DNA, connect

"–SH" to $5'$ end of DNA molecules, fix linear DNA on silicon surfaces to get DNA chips. Let $i = 1$;

2. Take out the marker probes i and respectively mix these marker probes with DNA chips sufficiently. Raise temperature to 94 °C and then fall slowly to 68 °C, do hybridization experiment, and take out all DNA chips and wash it with buffering solution of 68 °C. Then scan DNA chips with laser focusing microscope to detect the hybridization result and record fluorescence light dot of DNA chips.

3. Mix buffering solution of 94 °C with DNA chips sufficiently to react degeneration, and wash all marker probes on DNA chips with it;

4. If $i = n$, and return to 5; else let $i = i + 1$, and return to 2;

5. Analyze the record of each fluorescence light dot of DNA chips, and decode out all maximum weighted independent sets.

4 An Instance of the Maximum Weighted Independent Set Problem

We explain the effectiveness and correctness of DNA computing, using the maximum weighted independent set problem of simple undirected weighted graph $G = (V, E, W)$ which is shown in Fig. 3.

The DNA computing results are as follows.

1. Encode each vertex of the graph $G = (V, E, W)$, and $V = \{v_1, v_2, v_3, v_4, v_5\}$, $E = \{e_1, e_2, e_3, e_4, e_5, e_6\}$, $W = \{w_1, w_2, w_3, w_4, w_5\}$.

The chain length of encoding a_i $(i = 1, 2, \ldots, 5)$ is as follows. $|a_1| = 36byte$, $|a_2| = 72byte$, $|a_3| = 48byte$, $|a_4| = 60byte$, $|a_5| = 84byte$.

In accordance with the order of the vertex v_1, v_2, v_3, v_4, v_5 and 0 recognition site, select typeIIenzymes for the corresponding restriction endonuclease: $EcoR\,I$, Pst I, $Sal\,I$, $Kpn\,I$ and $BamH\,I$. The recognition sequences corresponding to these restriction endonuclease respectively are $5' - G_{\downarrow}AATTC - 3'$, $5' - C_{\downarrow}TGCAG - 3'$, $5' - G_{\downarrow}TCGAC - 3'$, $5' - G_{\downarrow}GTACC - 3'$, $5'- \quad G_{\downarrow}GATCC - 3'$. The positions of the arrow pointing respectively are the cutting sites of restriction enzymes, we recorded as $s_1 \downarrow t_1$, $s_2 \downarrow t_2$, $s_3 \downarrow t_3$, $s_4 \downarrow t_4$, $s_0 \downarrow t_0$.

Fig. 3 A weighted graph with *five vertices* (*Figures in brackets are the weights*)
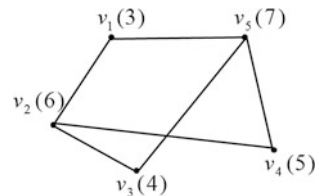
Synthesize double—stranded DNA of each vertex weight. Encodings of weight from a_1 to a_5 is: $a_1 : 5'- AATTC \cdots G - 3'$, $a_2 : 5' - TGCAG \cdots C - 3'$, $a_3 : 5' - TCGAC \cdots G - 3'$, $a_4 : 5' - GTACC \cdots G - 3'$, $a_5 : 5' - GATCC \cdots G - 3'$. The double—stranded DNA of each vertex weight is encoded in accordance with the chain length. Both ends of the DNA double strand are encoded with the recognition sequences corresponding to these restriction endonucleases. The ellipses indicate the random encoding of the four bases, but not including any recognition sequences of these five restriction enzymes.

2. In accordance with $s_0t_0s_1t_1s_2t_2s_3t_3s_4t_4$, use these five recognition sequences and synthesize closed circle DNA l^0, the chain length is $|l^0| = 30byte$, then insert foreign DNA double strand a_1, a_2, a_3, a_4, a_5 into closed circle DNA l^0. Synthesize closed circle DNA corresponding to the set of vertices of G $l^1 = (a_1a_2a_3a_4a_5)$, the chain length is $|l^1| = 30byte + |a_1| + |a_2| + |a_3| + |a_4| + |a_5| = 330byte$.

Making l^1 linearization Using restriction endonuclease *BamH I*, PCR amplification, and hybridization experiment, to obtain closed circle mixture l^1.

3. Do delete experiment. For edge $e_1 = v_1v_2 \in E$, divide the mixture into two equal parts, for each part do delete experiments respectively by the action of restriction endonuclease *EcoR I* and Pst *I*. After high performance liquid chromatography (HPLC), mix two parts sufficiently, the obtained mixture is $l^2 = \{(0a_2a_3a_4a_5), (a_10a_3a_4a_5)\}$;

For edge $e_2 = v_2v_3$, by the action of Pst *I* and *Sal I*, the obtained mixture is $l^3 = \{(00a_3a_4a_5), (a_10a_3a_4a_5), (0a_20a_4a_5), (a_100a_4a_5)\}$;
For edge $e_3 = v_2v_4$, by the action of Pst *I* and *Kpn I*, the obtained mixture is $l^4 = \{(00a_3a_4a_5), (a_10a_3a_4a_5), (000a_4a_5), (a_100a_4a_5), (00a_30a_5), (a_10a_30a_5), (0a_200a_5), (a_1000a_5)\}$;
For edge $e_4 = v_3v_5$, by the action of *Sal I* and *BamH I*, the obtained mixture is $l^5 = \{(000a_4a_5), (a_100a_4a_5), (0000a_5), (a_1000a_5), (0a_200a_5), (00a_3a_40), (a_10a_3a_40), (a_10a_3a_40), (000a_40), (a_100a_40), (00a_300), (a_10a_300), (0a_2000), (a_10000)\}$;
For edge $e_5 = v_4v_5$, by the action of *Kpn I* and *BamH I*, the obtained mixture is $l^6 = \{(0000a_5), (a_1000a_5), (0a_200a_5), (00a_300), (a_10a_300), (00000), (a_10000), (00a_300), (0a_2000), (000a_40), (a_100a_40), (00a_3a_40), (a_10a_3a_40)\}$;
For edge $e_6 = v_1v_5$, by the action of *EcoR I* and *BamH I I*, the obtained mixture is
$l^7 = \{(0000a_5), (0a_200a_5), (00a_300), (00000), (0a_2000), (000a_40), (00a_3a_40), (a_10000), (a_10a_300), (a_100a_40), (a_10a_3a_40)\}$.

4. Do gel electrophoresis experiment for the mixture, the result is:

$\{(00000)\} < \{(00a_300), (a_10000)\} < \{(000a_40)\} < \{(0a_2000)\} < \{(0000a_5), (a_10a_300)\} < \{(a_100a_40)\} < \{(00a_3a_40)\} < \{(a_10a_3a_40)\} < \{(0a_200a_5)\}$.

The longest subset is $\{(0a_200a_5)\}$.

5. Detect the mixture, and the result is $\{(0a_200a_5)\}$, the maximum weighted independent set is $\{(0a_200a_5)\}$.

5 Conclusion

In this paper, upon closed circle DNA computing model, an algorithm of the maximum weighted independent set problem is proposed. First encode and compound the initial closed circle DNA, then do delete experiments, obtain all independent sets and find the maximum weighted independent set using electrophoresis experiment, last detect and output results. Only using delete experiment, the algorithm has concise steps, simple idea and is ease to implement, which guarantees to reduce errors in solutions. Judging from the calculation process, the total operating number is $n + m$, n is the vertex-number of a graph and m is the edge-number.

But the algorithm requires a lot of recognition sequence and more DNA encoding, its scale is restricted by the limited restriction enzymes and the closed circle length. With the development of bioengineering technology, more and more restriction enzymes will be found, this problem will be changed. But, too long closed circle DNA will change their biochemical nature, DNA encoding problem is also an NP-complete problem, which affect further applications of this model. It would be our future work.

References

1. Adleman LM (1994) Molecular computation of solution to combinatorial problems. Science 266(11):1021–1024
2. Kang Z, Jin X (2006) Closed circle DNA algorithm of the minimal covering problem. Comput Eng Appl 42(20):7–9
3. Kang Z, Yanfang Y, Yuhua L, Lei Q (2007) Closed circle DNA model of weighted matching. J Huazhong Univ Sci Technol 35(8):63–66 (Nature Science Edition)
4. Kang Z, Yanfeng W, Wenbin L et al (2006) DNA algorithm of edge coloring problem on closed circle DNA. J Huazhong Univ Sci Technol 34(9):25–28 (Nature Science Edition)
5. Kang Z, Xiaojun T, Jin X (2007) Closed circle DNA algorithm of eight queens problem. Comput Eng Appl 43(6):4–6

6. Kang Z, Xiaojun T, Jin X (2007) An algorithm of assignment problem based on closed circle DNA. Comput Sci 34(12):211–213
7. You Y, Ziming D (2002) An algorithm for solving maximum independent set of a graph. Comput Dev Appl 15(6):13–14
8. Cui J, Yin Z, Yang J (2008) PNA-based DNA computing model for the maximum independent set problem. Int J Biomath 23(3):501–508
9. Yueke F, Xiaoli Q, Jin X (2010) Sticker model for maximum clique problem and maximum independent set. Chin J Comput 33(2):305–310
10. Kang Z, Xiaojun T, Wenbin L et al (2008) Closed circle DNA algorithm of the maximum independent set problem. Comput Eng 34(4):40–44
11. Xue W, Yi Z (2007) DNA solution of the maximum weighted independent set. J Electron Inf Technol 29(11):2693–2697

Bivariate Barycentric Rational Hermite Interpolaiton Based on the Lebesgue Constant Minimizing

Qianjin Zhao and Jie Qiao

Abstract Barycentric interpolation is considered to be the most stable formula for a rational Hermite interpolation. The core problem is to choose the optimal weights. In this paper, the optimal weights of the bivariate barycentric rational Hermite interpolation are obtained based on the Lebesgue constant minimizing. Then the solution of the optimization model can be obtained by the software LINGO. Further, the numerical examples are given to show the effectiveness of the new method.

Keywords Bivariate hermite interpolation · Barycentric · Lebesgue constant · Optimization · Weight

1 Introduction

The barycentric form of rational interpolation is a useful and powerful method in the interpolation. Recently, the barycentric rational interpolation and the barycentric rational Hermite interpolation is concerned a lot by many research scholars [1–6]. That's all because compared with other classical interpolation formulas the barycentric form of rational and Hermite interpolation offers a lot of advantages, such as numerical stability, no poles and no unattainable points [4].

In this paper, the Lebesgue constant of the bivariate barycentric rational Hermite interpolation has been constructed, and it is used to find out the optimal weights. Specifically, the weights as decision variables, the minimizing Lebesgue

Q. Zhao (✉) · J. Qiao
College of Science, Anhui University of Science and Technology, Huainan 232001, China
e-mail: qianjinzhao@yahoo.com.cn

J. Qiao
e-mail: qiaojie0302@163.com

constant of the bivariate barycentric rational Hermite interpolation is given for the objective function of optimization model. Then the solution of the optimization model can be obtained by the software LINGO. Finally, the numerical examples are given to show the effectiveness of the new method.

2 Univariate Barycentric Rational Hermite Interpolation

Let $r(x) \in R_{m,n}$, by $R_{m,n}$ the set of all functions with numerator degree $\leq m$ and denominator $\leq n$. Given $x_0, x_1, \ldots, x_N$ be $N + 1$ distinct nodes, the barycentric rational Hermite interpolation is presented as follows:

$$r(x) = \frac{\sum_{i=0}^{n} \sum_{k=0}^{s_i-1} \frac{w_{ik}}{(x-x_i)^{k+1}} \sum_{j=0}^{k} \frac{f_i^{(j)}}{j!}(x-x_i)^j}{\sum_{i=0}^{n} \sum_{k=0}^{s_i-1} \frac{w_{ik}}{(x-x_i)^{k+1}}} \tag{1}$$

for the data $(x_i, f_i^{(j)})$, $j = 0, 1, 2, \ldots, s_i-1$, $i = 0, 1, 2, \ldots, N$ where $x_i \neq x_j$ for $i \neq j$. In [4], the necessary conditions of barycentric rational Hermite interpolation which has no poles and no unattainable point occurs are given as follows:

$$sign\ w_{i,s_i-1} = (-1)^{s_i+1} sign\ w_{i+1,s_{i+1}-1}, \quad i = 0, 1, 2, \ldots, N-1 \tag{2}$$

$$w_{i,s_i-1} \neq 0, \quad i = 0, 1, 2, \ldots, N \tag{3}$$

3 Bivariate Barycentric Rational Hermite Interpolation

Following [7], the set of bivariate functions are considered as follows:

$$\Pi^{nm} = \{(x_i, y_j) | i = 0, 1, \ldots, n; j = 0, 1, \ldots m\}, D = (a, b) \times (c, d),$$

where $\Pi^n = \{x_i | i = 0, 1, \ldots, n\} \subset (a, b), x_0 < x_1 < \ldots < x_n$, and $\Pi^m = \{y_j | j = 0, 1, \ldots, m\} \subset (c, d)$, and let a sufficiently smooth bivariate function $f(x, y)$ be defined in D, the set of interpolation data of Hermite type given by $f_{ij}^{(r,s)} = \frac{\partial^{r+s}}{\partial x^r s^s} f(x_i, y_j)$, $(i = 0, 1, \ldots, n; j = 0, 1, \ldots, m)$, $(r = 0, 1, \ldots, \alpha_i - 1; s = 0, 1, \ldots, \beta_j - 1)$.

To obtain the bivariate barycentric rational Hermite interpolation, the rational function is considered as follows:

$$R(x,y) = \frac{\sum_{i=0}^{n}\sum_{k=0}^{\alpha_i-1}\frac{w_{ik}}{(x-x_i)^{k+1}}\sum_{l=0}^{k}\frac{\partial^l r_i(y)}{l!\partial x^l}(x-x_i)^l}{\sum_{i=0}^{n}\sum_{k=0}^{\alpha_i-1}\frac{w_{ik}}{(x-x_i)^{k+1}}} = \frac{P(x,y)}{Q(x,y)}, \tag{4}$$

including:

1.

$$r_i(y) = \frac{\sum_{j=0}^{m}\sum_{h=0}^{\beta_j-1}\frac{u_{jh}}{(y-y_j)^{h+1}}\sum_{t=0}^{h}\frac{\partial^t f(x_i,y_j)}{t!\partial y^t}(y-y_j)^t}{\sum_{j=0}^{m}\sum_{h=0}^{\beta_j-1}\frac{u_{jh}}{(y-y_j)^{h+1}}}, \qquad (i=0,1,\ldots,n).$$

2. Let $n+1$ distinct points $x_0 < x_1 < \ldots < x_n$ be given, together with corresponding weights w_{ik} $(i=0,1,\ldots,n;\ k=0,1,\ldots,\alpha_i-1)$. At the same time, let $y_0 < y_1 < \ldots < y_m$ be given, together with corresponding weights u_{jh} $(j=0,1,\ldots,m;\ h=0,1,\ldots,\beta_j-1)$. The weights meet the conditions as follows:

$$sign(w_{i,\alpha_i-1}) = (-1)^{\alpha_{i+1}}sign(w_{i+1,\alpha_{i+1}-1}), \quad (i=0,1,\ldots,n-1), \tag{5}$$

$$sign\left(u_{j,\beta_j-1}\right) = (-1)^{\beta_{j+1}}sign\left(u_{j+1,\beta_{j+1}-1}\right), \quad (j=0,1,\ldots,m-1). \tag{6}$$

3. From [4], the weights of the bivariate barycentric rational interpolation should satisfy as follows:

$$q(x) = \omega(x)\sum_{i=0}^{n}\sum_{k=0}^{\alpha_i-1}\frac{w_{ik}}{(x-x_i)^{k+1}} > 0,\ \omega(x) = \prod_{i=0}^{n}(x-x_i)^{\alpha_i}, \tag{7}$$

$$q(y) = \tilde{\omega}(y)\sum_{j=0}^{m}\sum_{h=0}^{\beta_j-1}\frac{u_{jh}}{(y-y_j)^{h+1}} > 0,\ \tilde{\omega}(y) = \prod_{j=0}^{m}(y-y_j)^{\beta_j}, \tag{8}$$

$$\left|\frac{\omega(x)}{q(x)}\right| \le 1,\ \left|\frac{\tilde{\omega}(y)}{q(y)}\right| \le 1. \tag{9}$$

4. To satisfy interpolation conditions: $R^{(r,s)}(x_i,y_j) = f_{ij}^{(r,s)}$, $\forall(x_i,y_j) \in \Pi^{nm}$, all the weights should meet the conditions:

$$w_{ik} \ne 0,\ u_{jh} \ne 0. \tag{10}$$

3.1 The Lebesgue Constant of the Bivatiate Barycentric Rational Hermite Interpolation

The Lebesgue constant has been studied intensively in polynomial interpolation [8, 9]. However, from [10] the optimal weights can then be obtained by the Lebesgue constant minimizing. Basis on this idea let us construct the Lebesgue constant of the bivariate barycentric rational Hermite interpolation:

$$
\Lambda_{nm} = \max_{a<x<b,c<y<d} \sum_{i=0}^{n} \sum_{j=0}^{m} \left| \frac{\sum_{k=0}^{\alpha_i-1} \sum_{h=0}^{\beta_j-1} \sum_{l=0}^{k} \sum_{t=0}^{h} \frac{w_{ik}u_{jh}}{(x-x_i)^{k+1-l}(y-y_j)^{h+1-t}}}{\sum_{i=0}^{n} \sum_{j=0}^{m} \sum_{k=0}^{\alpha_i-1} \sum_{h=0}^{\beta_j-1} \frac{w_{ik}u_{jh}}{(x-x_i)^{k+1}(y-y_j)^{h+1}}} \right|, \tag{11}
$$

so that the minimizing Lebesgue constant of the bivariate barycentric rational Hermite interpolation is:

$$
\min \Lambda_{nm} = \min_{a<x<b,c<y<d} \max_{a<x<b,c<y<d} \sum_{i=0}^{n} \sum_{j=0}^{m} \left| \frac{\sum_{k=0}^{\alpha_i-1} \sum_{h=0}^{\beta_j-1} \sum_{l=0}^{k} \sum_{t=0}^{h} \frac{w_{ik}u_{jh}}{(x-x_i)^{k+1-l}(y-y_j)^{h+1-t}}}{\sum_{i=0}^{n} \sum_{j=0}^{m} \sum_{k=0}^{\alpha_i-1} \sum_{h=0}^{\beta_j-1} \frac{w_{ik}u_{jh}}{(x-x_i)^{k+1}(y-y_j)^{h+1}}} \right|.
$$
$$
\tag{12}
$$

3.2 Optimization Model for the Optimal Weights

Through the above analysis, to ensure that the bivariate barycentric rational Hermite function guaranteed the interpolation conditions and has no pole or no unattainable point, we can see that the optimization model involves the following several:

$$
\max_{a<x<b,c<y<d} \sum_{i=0}^{n} \sum_{j=0}^{m} \left| \frac{\sum_{k=0}^{\alpha_i-1} \sum_{h=0}^{\beta_j-1} \sum_{l=0}^{k} \sum_{t=0}^{h} \frac{w_{ik}u_{jh}}{(x-x_i)^{k+1-l}(y-y_j)^{h+1-t}}}{\sum_{i=0}^{n} \sum_{j=0}^{m} \sum_{k=0}^{\alpha_i-1} \sum_{h=0}^{\beta_j-1} \frac{w_{ik}u_{jh}}{(x-x_i)^{k+1}(y-y_j)^{h+1}}} \right| = \min,
$$

$s.t. \quad w_{i,s_i-1} \neq 0, \ u_{jh} \neq 0, \quad (i=0,1,2,\ldots,n; \ j=0,1,\ldots,m),$

$sign(w_{i,\alpha_i-1}) = (-1)^{\alpha_i+1} sign(w_{i+1,\alpha_{i+1}-1}), \quad (i=0,1,\ldots,n-1),$

$sign(u_{j,\beta_j-1}) = (-1)^{\beta_j+1} sign(u_{j+1,\beta_{j+1}-1}), \quad (j=0,1,\ldots,m-1),$

$$q(x) = \omega(x) \sum_{i=0}^{n} \sum_{k=0}^{\alpha_i-1} \frac{w_{ik}}{(x-x_i)^{k+1}} > 0, \ \omega(x) = \prod_{i=0}^{n} (x-x_i)^{\alpha_i},$$

$$q(y) = \tilde{\omega}(y) \sum_{j=0}^{m} \sum_{h=0}^{\beta_j-1} \frac{u_{jh}}{(y-y_j)^{h+1}} > 0, \ \tilde{\omega}(y) = \prod_{j=0}^{m} (y-y_j)^{\beta_j},$$

$$\left|\frac{\omega(x)}{q(x)}\right| \leq 1, \ \left|\frac{\tilde{\omega}(y)}{q(y)}\right| \leq 1.$$

Moreover, $\sum_{i=0}^{n} \sum_{k=0}^{\alpha_i-1} |w_{ik}| = 1$ and $\sum_{j=0}^{m} \sum_{h=0}^{\beta_j-1} |u_{jh}| = 1$ are added to the optimization model for ensuring the uniqueness of the optimal solutions.

4 Numerical Examples

To illustrate the feasibility, we present a simple example in this section. Given $f(x, y) = \frac{1}{1+x^2} + \frac{y}{1+y^2}$ as the interpolated, and the set of the points $\{(x_i, y_j) = (i, j) | \ i = 0, 1, \ldots, n; \ j = 0, 1, \ldots, m\} \subset \Pi_{xy}^{22}$, the corresponding values as follows (Tables 1, 2, 3):

Table 1 The values of $f(x_i, y_j)$

	$x_0 = 0$	$x_1 = 0.5$	$x_2 = 1$
$y_0 = 0$	1.00	1.40	1.50
$y_1 = 0.5$	0.80	1.20	1.30
$y_2 = 1$	0.50	0.90	1.00

Table 2 The values of $\frac{\partial f(x_i, y_j)}{\partial x}$

	$x_0 = 0$	$x_1 = 0.5$	$x_2 = 1$
$y_0 = 0$	0	0	0
$y_1 = 0.5$	−0.64	−0.64	−0.64
$y_2 = 1$	−0.50	−0.50	−0.50

Table 3 The values of $\frac{\partial f(x_i, y_j)}{\partial y}$

	$x_0 = 0$	$x_1 = 0.5$	$x_2 = 1$
$y_0 = 0$	1.00	0.48	0
$y_1 = 0.5$	1.00	0.48	0
$y_2 = 1$	1.00	0.48	0

From the function $f(x, y) = \frac{1}{1+x^2} + \frac{y}{1+y^2}$, it is easy to know that the values of $\frac{\partial^2 f(x_i, y_j)}{\partial x \partial y} = 0$, $(i = 0, 1, 2, \ldots, n; j = 0, 1, \ldots, m)$.

Then the solution of the optimization model can be obtained by the software LINGO. We get the optimal weights as follows:

$w_{00} = 0.1380974$, $w_{01} = 0.1000000e - 5$, $w_{10} = 0.2113745e - 5$,
$w_{11} = 0.1000000e - 5$, $w_{20} = 0.5908288e - 1$, $w_{21} = 0.8028156$,
$u_{00} = 0.1000000e - 5$, $u_{01} = 0.2557674$, $u_{10} = 0.2530650$,
$u_{11} = 0.4108120$, $u_{20} = 0.1000000e - 5$, $u_{21} = 0.8035355e - 1$.

In order to show the effectiveness of the new method, the function interpolated, the interpolant and the error function are drawn by using the software MATLAB as follows (Figs. 1, 2, 3):

Fig. 1 The interpolated function
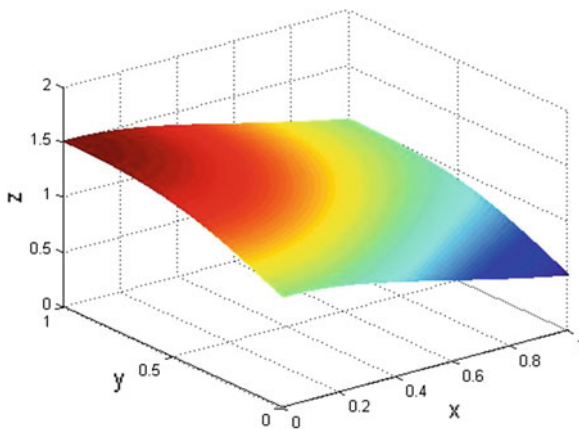
Fig. 2 The interpolant
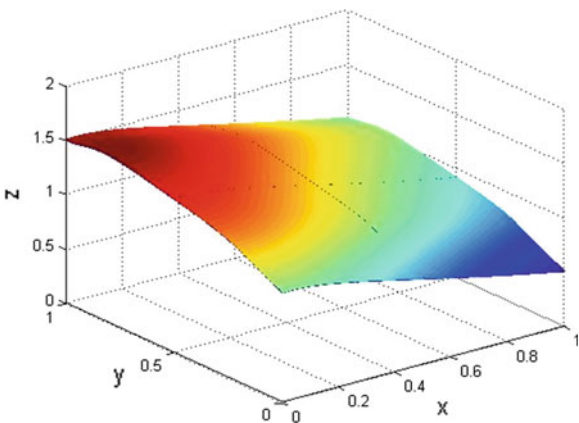
Fig. 3 The error function
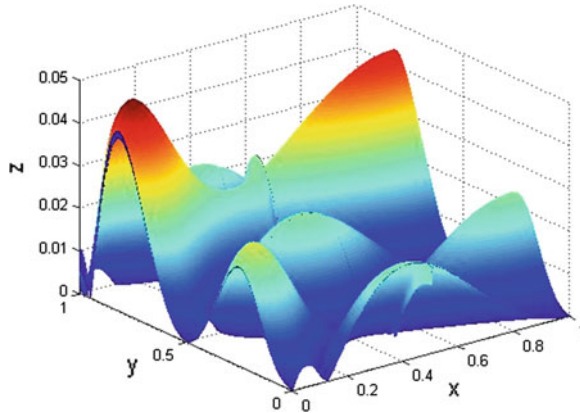
5 Conclusion

We have introduced a uniform method for the bivariate barycentric rational Hermite interpolation based on the Lebesgue constant minimum, and the Lebesgue constant of the bivariate barycentric rational Hermite interpolation has been constructed. Properties of the above-mentioned:

- Numerical stability.
- Without poles and unattainable points.
- Small calculation quantity.
- Arbitrarily high approximation orders.

References

1. Floater MS, Kai H (2007) Barycentric rational interpolation with no poles and high rates of approximation. Numer Math 107:315–331
2. Klein G, Berrut JP (2012) Linear rational finite differences from derivatives of barycentric rational interpolants. SIAM J Numer Anal 50:643–656
3. Bos L, DeMarchi S, Kai H (2011) On the lebesgue constant of Berrut's rational interpolant at equidistant nodes. Comput Appl Math 236:504–510
4. Claus S, Wilhelm M, Kiinelsau W (1991) Hermite interpolation: the Barycentric approach. Computing Special issue on archives for informatics and numerical computation 46:35–51
5. Berriochoa E, Cachafeiro A (2010) Algorithms for solving hermite interpolation problems using the fast Fourier transform. J Comput Appl Math 235:882–894

6. Ana M, Jos′e-Javier M (2009) Unique solvability in bivariate hermite interpolation. Electron Trans Numer Anal 34:20–30
7. Gasca M, Mart′i Nez JJ (1990) On the solvability of bivariate Hermite–Birkhoff interpolation problems. J Comput Appl Math 32:77–82
8. Damelin SB (1998) The Lebesgue function and Lebesgue constant of Lagrange interpolation for Erdos weights. Approximation Theory 94:235–262
9. Mills TM, Smith SJ (1992) The Lebesgue constant for Lagrange interpolation on equidistant nodes. Numer Math 61:111–115
10. Berrut JP (1997) Lebesgue constant minimizing linear rational interpolation of continuous functions over the interval. Comput Math Appl 33:77–86

Efficiency of Parallel Computing Based on .Net 4.5

Wei Hao, Rongrong Gu, Kelei Sun and Ping Ren

Abstract In recent decades, with the continuous development of the CPU, the speed of single-core processors is increasingly close to the limit. Therefore, multi-core processors, through the internal integration of multi cores, solve the bottleneck in the development of the CPU. Over the past decade, the pace of development of multi-core technology is very fast, but the development of the software has not kept up with the pace of hardware. Most software and program failed to fully utilize the performance of multi-core processors. This paper introduces the parallel programming knowledge of Microsoft .Net 4.5 through the design of an experiment seeking the largest prime number demonstrates parallel the implementation process. Meanwhile, through the comparison of the time consumption between the parallel computation and traditional method, the author draw a conclusion that the computation efficiency is greatly improved by parallel computation then the traditional method in case of large amount of data.

Keywords Parallel computing · Comparison · .Net 4.5 · Multi-core processor

1 Introduction

Moore's law is the observation that over the history of computing hardware, the number of transistors on integrated circuits doubles approximately every 2 years. The period often quoted as "18 months" is due to Intel executive David House, who predicted that period for a doubling in chip performance (being a combination of the effect of more transistors and their being faster) [1].

W. Hao (✉) · R. Gu · K. Sun · P. Ren
School of Computer Science and Engineering, Anhui University of Science and Technology,
Huainan 232001, China
e-mail: whao@aust.edu.cn

However, the free lunch is not always provided, the CPU speed cannot be unlimitly upgrade. The existing CPU technology is limited by the laws of physics. When the transistor is downsized to certain scale, due to very slight distance between the control of the current of the transistor gate and the gate oxide, the electron drift occurs. If this occurs, the transistor will lost reliability, because the transistor under such conditions cannot effectively control the movement of electrons, thereby making it difficult to let it be able to maintain a stable "1" or "0" state [2].

In order to break through these physical limitations and continue to enhance the computing power of the CPU, the CPU designers take another way: multi-core processing. Multi-core processing integrates multiple CPU processing cores into a single CPU, called "multi-core CPU". Thus, each CPU processing core doesn't increase computing power, but reasonable task assigned by calculation to run programs on multiple CPU processing cores, we can achieve the same purpose to speed up the processing performance of the CPU [3].

CPU's performance is increased by way of multi-core and can only be fully utilized by parallel computing. If we really want to fully utilize the performance of multi-core processors, it must rely on a well-designed software which can be divided into multiple tasks and be executed in parallel subtasks in support of the operating system (or specific software operating platform). The development of parallel computing software system is much more difficult than the traditional program design and presents new challenges for software engineers [4].

This paper is to test the efficiency improvement of .Net 4.5 by designing an experiment of Big Prime Finder seeking the largest prime number.

2 The Basic Principle of Parallel Computing

This section introduces the related basic concepts of parallel processing.

2.1 Multi-Core Process

A multi-core processor is a single computing component with two or more independent actual central processing units (called "cores"), which are the units that read and execute program instructions [5].

Figure 1 demonstrates the structure of multi-core process. CPU is composed of executable core, Cache, and external system bus interface. Executable core is a collection of available functional units which is used to calculate and control in the microprocessor. Cache is a small area of fast memory used by CPU. It provides data for CPU and exchange data with system bus through external system bus interface which is used as a bridge between CPU Cache and system bus.

Fig. 1 Multi-core process
structure
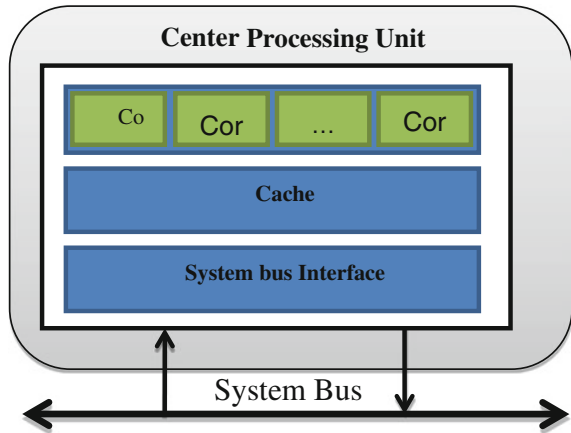
2.2 Parallelism

Parallelism refers to the simultaneous executions of multiple tasks on a multi-core
CPU (or more single-core CPU [6]). In the process of running these tasks, there are
always at least two or more tasks simultaneously running at any point in time.
Figure 2 demonstrates the principle of Parallelism.

2.3 Concurrency

Parallelism refers to the simultaneous executions of multiple tasks on a multi-core
CPU (or more single-core CPU [7]). In the process of running these tasks, there are
always at least two or more tasks simultaneously running at any point in time.
Figure 2 demonstrates the principle of Parallelism (Fig. 3).
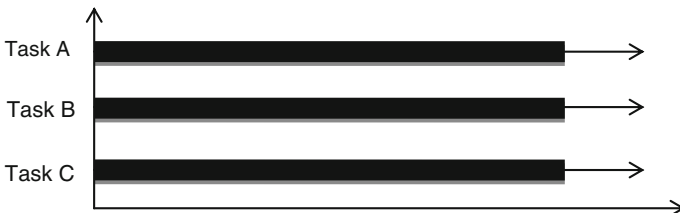
Fig. 2 Parallelism schematic diagram

Fig. 3 Parallelism schematic diagram

3 Introduction to .Net 4.5 Parallel Computing

Microsoft has kept in step with the computing industry's quest for Parallelism. Since .Net 4.0, Parallelism has been an addition to .NET because CPU clock speeds have stagnated and manufacturers have shifted their focus to increasing core counts. In 2012, Microsoft published the latest version of VS, Visual Studio 2012 with .Net 4.5 which has some new feature for parallelism [8].

.NET Parallelism is divided into three parts: Parallel LINQ, the Parallel class and the task parallelism constructs. The concurrent collections are new to Framework 4.5 and are collectively known as PFX (Parallel Framework). The Parallel class together with the task parallelism constructs is called the Task Parallel Library or TPL [9].

This is problematic for us as programmers because traditional standard single-threaded code will not automatically run faster as a result of those extra cores. That being the case, this paper will take a look at parallel programming as done in the C# 4.5 language and the .NET 4.5 runtime [10].

This paper is to compare the efficiency between traditional programming and the parallel programming. Parallelism is used to improve the performance of computing, therefore, how much does it change is my main topic.

4 Experiments

In order to find out the parallel computational efficiency improvements of .Net 4.5 parallelism, the paper designs an evaluation experiment of the largest prime number. In this experiment, we check each number from 3 to n, to determine whether it is a prime number, and then records the time it takes. We used two methods, one is the traditional iteration, and another is to use. Net 4.5 parallelism [11].

4.1 Create PrimeFinder Class for .Net 4.5 Parallelism Calculation

First, the PrimeFinder class is created as below.

```
class PrimeFinder{
    public PrimeFinder() { }
    public PrimeFinder(long t) { target = t; }
}
```

It has two constructors, one is default, and another with an input parameter is used to get input which determines the max number to check.

Now we need to find largest prime number, then we must begin to establish a function for determining whether a given number n is a prime number. The function is created as below.

```
//determine whether n is a prime number
public bool IsPrime(long n) {
    double end = Math.Sqrt(n);
    for (long i = 3; i <= end; i++){
        if (n % i == 0) {
            return false;
        }
    }
        return true;
    }
```

And then it defines some variants for calculation.

```
private long n = 3;             //current number to be
check
private long maxprime = 3;      //max prime found
private long target = 1000000;//max number to be check
//10 tasks is defined to finish the job
private Task<long>[] tasks = new Task<long>[10];
```

The algorithm below is the main part of the calculation. It records time before calculation, and outputs total time cost after the main calculation is done. The main calculation part contains 10 tasks witch calculate simultaneously. Each task gets variable n from the class, and then plus 1 to n. After obtaining n, the task begins to calculate whether n is a prime number. If n is a prime, the task returns n, otherwise zero. When a task is complete, it gets n again from the class until the loop ends.

```
private void Find(){
    DateTime dt = DateTime.Now;




    Task task = new Task(() => {
            Console.WriteLine(dt + ": Mission Start.");
            while (running) {
                if (n > target) {
                    TimeSpan ts = DateTime.Now - dt;
                    Console.WriteLine(DateTime.Now + ":
Mission Complete. Total Second: " + ts.TotalSeconds);
                    break;
                }
                for (int i = 0; i < tasks.Length; i++)
{
                    if (tasks[i].Status == TaskSta-
tus.RanToCompletion) {
                        Check(i);
                    }
                }
            }
        });
    task.Start();}
Every loop starts in a task.
    public void Check(int i) {
        tasks[i] = Task<long>.Factory.StartNew((obj)
=>{
            long num = (long)obj;
            return IsPrime(n) ? num : 0; }, n);
        n += 2;
        if (tasks[i].Result > 0)  maxprime = n;
 }
```

4.2 Traditional Method

Traditional method is very easy. A for loop is used to find the max prime in given number and records the time it takes.

```
public long Tradition(long target){
    DateTime dt = DateTime.Now;
    Console.WriteLine(dt + ": Mission Start.");
    long max = 3;
    for (int i = 0; i < target; i+= 2) {
        if (IsPrime(i))
            max = i; }
    TimeSpan ts = DateTime.Now - dt;
Console.WriteLine(DateTime.Now + ": Mission Complete.
Total Second: " + ts.TotalSeconds);
return max;
}
```

Table 1 Testing environment parameters

Items	Parameters
CPU	i3 2120
Memory	2G
Display card	HD 6550
OS	Win7 64bit sp1

Table 2 Test result

	10 K (s)	100 K (s)	1,000 K (s)	10,000 K (s)
.Net 4.5 parallelism	0.1	0.6	7	86
Traditional method	0.05	0.7	9	123

Fig. 4 Result
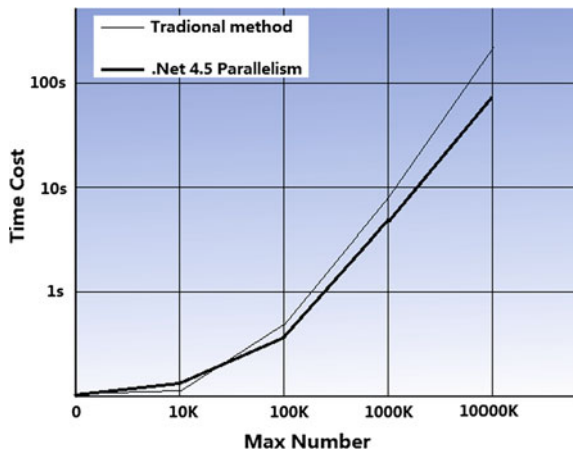
5 Results

After the code is finished, the test begins. This experiment takes in the following machine (Table 1).

4 group tests are carried out. Each test has been done for 3 times and then averaged. The data collected is in the Table 2 (Fig. 4).

6 Conclusions

According to experimental results, it can be clearly seen that the efficiency of the parallel computing in small amount of computation takes more time than traditional method. The reason is that parallel computing takes more extra time to manage tasks, therefore, more time is taken than the traditional method, so the efficiency is reduced. However, with the increase of amount of computation, the parallel computing efficiency is higher, and it can significantly improve the computational efficiency. So in that case, it is strongly recommended that the calculated using parallel manner.

References

1. Parallel Computing Concepts via C# 4.0 logicchild. http://www.codeproject.com/Articles/71369/Parallel-Computing-Concepts-via-C-4-0. 4 Nov 2010
2. Patterns of Parallel Programming, by Microsoft's Stephen Toub
3. Chen G (1999) Parallel computing: architecture algorithm programming. Beijing Advanced Education Press
4. Isard M et al (2007) Dryad: distributed data-parallel programs from sequential building blocks. pp 59–72
5. Chen G (1988) A partitioning selection algorithm on multiprocessors. J Comput Sci Technol 3:241–250
6. Chen G et al (2008) Proceedings of the inaugural symposium on parallel algorithms, architectures and programming, 16–18 September 2008
7. Chen Y, Lin P, Bao Y, He Y (2009) Application of the distributed and parallel computation in spectroscopy signal processing. Spectrosc Spectral Anal 29(4)
8. Yu R, Wu L, Qu S (2011) Parallel computing and visualization method of radar net's detection ability. Syst Eng Electron 33(11)
9. Li N, Chen Z, Gong G, Peng X (2010) Application of multi-core parallel computing technology in scene matching simulation. Syst Eng Electron 32(2)
10. Fan X, Wu R (2009) Parallel computing of clamp structure in tahoe frame. Chin J Comput Phys 26(5)
11. Wang Z, Hu H (2011) Parallel computation techniques for dynamic description logics reasoning. J Comput Res Dev 48(12)

Boundary Value Problem for a Coupled System of Nonlinear Fractional Differential Equation

Ya-ling Li and Shi-you Lin

Abstract This paper is concentrated on the following coupled system of the nonlinear fractional differential equation

$$
\begin{cases}
D^{\alpha}u(t) = f(t, v(t)) + \displaystyle\int_0^t K(s, v(s))ds, & 5 < \alpha, \beta \leq 6, \ 0 < t < 1 \\
D^{\beta}v(t) = g(t, u(t)) + \displaystyle\int_0^t H(s, u(s))ds \\
u(1) = \lim_{t \to 0} u(t) \cdot t^{2-\alpha} = v(1) = \lim_{t \to 0} v(t) \cdot t^{2-\beta} = 0.
\end{cases}
$$

where f, K, g, $H : [0, 1] \times \Re \rightarrow [0, +\infty)$ are the positive continuous functions. D^{α} and D^{β} are the standard Riemann–Liouville fractional derivatives with the order α, β, respectively. We give the existence and the uniqueness of the solution by using the Schauder fixed point theorem and the generalized Gronwall inequality.

Keywords Coupled system · Riemann–Liouville fractional derivative · Schauder fixed point theorem · Fractional differential equation

Mathematics Subject Classification (2000) 26A33 34A08

1 Introduction

It is well known that the theory of fractional calculus plays an important role in mathematics. It is extensively applied to the biological realm. Therefore, the boundary value problems for the fractional differential equations have attracted more and more attention in recent years.

Y. Li · S. Lin (✉)
School of Mathematics and Statistics, Hainan Normal University, Haikou, Hainan 571158, People's Republic of China
e-mail: linsy1111@yahoo.cn

In this paper, we will study the following boundary value problem

$$
\begin{cases}
D^\alpha u(t) = f(t, v(t)) + \int_0^t K(s, v(s))ds, & 5 < \alpha, \beta \le 6, \ 0 < t < 1 \\
D^\beta v(t) = g(t, u(t)) + \int_0^t H(s, u(s))ds \\
u(1) = \lim_{t \to 0} u(t) \cdot t^{2-\alpha} = v(1) = \lim_{t \to 0} v(t) \cdot t^{2-\beta} = 0.
\end{cases}
\tag{1}
$$

Here, f, K, g, $H : [0, 1] \times \Re \to [0, +\infty)$ are all the positive continuous functions. D^α and D^β are the fractional derivatives with the Riemann–Liouville type which is defined as below

Definition 1 [1–4] The αth Riemann–Liouville fractional derivative of a function $u : [0, +\infty) \mapsto \Re$ is defined by

$$
D^\alpha u(t) := \left(\frac{d}{dt}\right)^n (I^{n-\alpha} u)(t) = \frac{1}{\Gamma(n-\alpha)} \left(\frac{d}{dt}\right)^n \int_0^t (t-s)^{n-\alpha-1} u(s)ds,
$$

where $\alpha > 0$, $n = [\alpha] + 1$, $[\alpha]$ is the largest integer which is less than α. And $I^{n-\alpha}$ denotes the $(n - \alpha)$th Riemann–Liouville fractional integral.

We first present a lemma that will be used in the proof of our main result:

Lemma 1 [1–4] *Let* $\alpha > 0$, *then*

$$
I^\alpha D^\alpha u(t) = u(t) + \sum_{i=1}^n c_i t^{\alpha-i},
$$

where c_i, $i = 1, 2, \ldots, n$ *are some constants in* $\Re$ *and* $n = [\alpha] + 1$.

Now we state the main result as follows.

2 Main Results

Theorem 1 *If there exists a constant A satisfying*

$$
\begin{cases}
\frac{1}{\Gamma(\alpha+1)} \max_{0 \le s \le 1, |t| \le A} f(s, t) + \frac{1}{\Gamma(\alpha+2)} \max_{0 \le s \le 1, |t| \le A} K(s, t) \le A; \\
\frac{1}{\Gamma(\beta+1)} \max_{0 \le s \le 1, |t| \le A} g(s, t) + \frac{1}{\Gamma(\beta+2)} \max_{0 \le s \le 1, |t| \le A} H(s, t) \le A.
\end{cases}
$$

Then the boundary value problem (1) *has a solution at least. In addition, if*

$$
I^\alpha D^\alpha u(1) = C_\alpha, \qquad I^\beta D^\beta v(1) = C_\beta,
\tag{2}
$$

where C_α, C_β *are two constants. And for any* $s \in [0, 1], u, v \in \Re$,

$$\begin{cases} |f(s,\,u)-f(s,\,v)| \le f_1(|u-v|); \\ |K(s,\,u)-K(s,\,v)| \le f_2(|u-v|); \\ |g(s,\,u)-g(s,\,v)| \le f_3(|u-v|); \\ |H(s,\,u)-H(s,\,v)| \le f_4(|u-v|). \end{cases} \qquad (3)$$

Here $F_1 = f_1 + f_2 : [0,+\infty) \to [0,+\infty)$ *and* $F_2 = f_3 + f_4 : [0,+\infty) \to [0,+\infty)$ *are two strictly monotonic increasing functions which satisfy that*

$$(F_1 + F_2)(0) = 0, \quad \int_0^1 \frac{1}{(F_1+F_2)(s)}\,ds = +\infty,$$

Then the solution of the problem (1) *is unique.*

3 Proof of Main Result

Let

$$\begin{cases} h_v(t) = f(t,\,v(t)) + \int_0^t K(s,\,v(s))ds; \\ j_u(t) = g(t,\,u(t)) + \int_0^t H(s,\,u(s))ds. \end{cases}$$

We first prove the following lemma

Lemma 2 *The solution of* (1) *can be described as below*

$$\begin{cases} u(t) = \int_0^1 G_\alpha(t,\,s)h_v(s)ds \\ v(t) = \int_0^1 G_\beta(t,\,s)j_u(s)ds. \end{cases} \qquad (4)$$

where

$$G_\alpha(t,\,s) = \begin{cases} \frac{(t-s)^{\alpha-1}-t^{\alpha-1}(1-s)^{\alpha-1}}{\Gamma(\alpha)}, & 0 \le s \le t \le 1 \\ -\frac{t^{\alpha-1}(1-s)^{\alpha-1}}{\Gamma(\alpha)}, & 0 \le t \le s \le 1 \end{cases}$$

$$G_\beta(t,\,s) = \begin{cases} \frac{(t-s)^{\beta-1}-t^{\beta-1}(1-s)^{\beta-1}}{\Gamma(\beta)}, & 0 \le s \le t \le 1 \\ -\frac{t^{\beta-1}(1-s)^{\beta-1}}{\Gamma(\beta)}, & 0 \le t \le s \le 1. \end{cases}$$

Proof From (1), $D^\alpha u(t) = h_v(t)$, which implies that

$$I^\alpha D^\alpha u(t) = I^\alpha h_v(t).$$

By **Lemma 1**,

$$u(t) = I^\alpha h_v(t) + \sum_{i=1}^6 C_i t^{\alpha-i}.$$

Since $\lim_{t\to 0} u(t) \cdot t^{2-\alpha} = 0$, we have

$$u(t) = I^\alpha h_v(t) + C_1 t^{\alpha-1}.$$

This, combining with the that $u(1) = 0$, yields

$$C_1 = -\frac{1}{\Gamma(\alpha)} \int_0^1 (1-s)^{\alpha-1} h_v(s)ds = -I^\alpha D^\alpha u(1)$$

Therefore,

$$u(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} h_v(s)ds - \frac{1}{\Gamma(\alpha)} \int_0^1 t^{\alpha-1}(1-s)^{\alpha-1} h_v(s)ds$$

$$= \int_0^1 G_\alpha(t,s) h_v(s)ds,$$

where

$$G_\alpha(t,\,s) = \begin{cases} \frac{(t-s)^{\alpha-1} - t^{\alpha-1}(1-s)^{\alpha-1}}{\Gamma(\alpha)}, & 0 \le s \le t \le 1 \\ -\frac{t^{\alpha-1}(1-s)^{\alpha-1}}{\Gamma(\alpha)}, & 0 \le t \le s \le 1 \end{cases}$$

Similarly, we can get $v(t) = \int_0^1 G_\beta(t,s) j_u(s)ds$. Here

$$G_\beta(t,s) = \begin{cases} (t-s)^{\beta-1} - t^{\beta-1}(1-s)^{\beta-1} & 0 \le s \le t \le 1 \\ -\frac{t^{\beta-1}(1-s)^{\beta-1}}{\Gamma(\beta)}, & 0 \le t \le s \le 1 \end{cases}$$

This completes the proof of **Lemma 2**.

Proof of Theorem 1 Let $X = C[0,1], \|u\| = \max_{t\in[0,1]} |u(t)|$ and $\|(u,v)\| = \max_{t\in[0,1]} \{\|u\|, \|v\|\}$. It is easy to see that $(X \times X, \|(u,v)\|)$ is a Banach space.

For any $(u,v) \in X \times X$, we suppose that

$$T(u,v)(t) = \left(-\int_0^1 G_\alpha(t,s) \left[f(s,-v(s)) + \int_0^s K(\eta, -v(\eta))d\eta \right] ds, \right.$$

$$\left. -\int_0^1 G_\beta(t,s) \left[g(s,-u(s)) + \int_0^s H(\eta, -u(\eta))d\eta \right] ds \right)$$

$$= (T_1 v(t), T_2 u(t)).$$

Define

$$U = \{ (u(t), v(t)) | 0 < u(t), v(t) \in X, \|(u(t), v(t))\| \le A, \quad t \in [0,1] \},$$

then U is a bounded closed set of $X \times X$. We proceed in two steps to prove that $T : U \to U$ is a completely continuous operator.

Step 1. $T : U \to U$ is compact.

Firstly, it is easy to see that:

1. $G_\alpha(t, s)$ and $G_\beta(t, s)$ are two continuous functions and $G_\alpha(t, s) \leq 0$, $G_\beta(t, s) \leq 0$;
2. $G_\alpha(t, s) \leq \frac{1}{\Gamma(\alpha)}(1-s)^{\alpha-1}$, $G_\beta(t, s) \leq \frac{1}{\Gamma(\beta)}(1-s)^{\beta-1}$.

For any $(u, v) \in U$, we have

$$0 \leq T_1 v(t) = -\int_0^1 G_\alpha(t, s)\left[f(s, -v(s)) + \int_0^s K(\eta, -v(\eta))d\eta\right] ds$$

$$\leq \frac{1}{\Gamma(\alpha)}\int_0^1 (1-s)^{\alpha-1}\left[\max_{0 \leq s \leq 1, |t| \leq A} f(s, t) + s \cdot \max_{0 \leq s \leq 1, |t| \leq A} K(s, t)\right] ds \leq A.$$

Similarly, $0 < T_2 u(t) \leq A$. Therefore, $\|T(u, v)\| \leq A$. $T : U \to U$ and the functions in TU are uniform bounded. In addition, by the continuity $G_\alpha(t, s)$ and $G_\beta(t, s)$ on $[0, 1] \times [0, 1]$, these functions are also uniform continuity on $[0, 1] \times [0, 1]$. So TU is equi-continuous. By Ascoli-Arzela theorem [5], $T : U \to U$ is a compact operator.

Step 2. $T : U \to U$ is continuous.

For any $(u_n, v_n), (u_0, v_0) \in U$, $\|(u_n, v_n) - (u_0, v_0)\| \to 0$ $(n \to \infty)$. Since f, K are the continuous functions, we have

$$\|T_1 v_n(t) - T_1 v_0(t)\| = \left\|\int_0^1 G_\alpha(t, s)\{[f(s, -v_n(s)) - f(s, -v_0(s))]\right.$$

$$\left. + \left[\int_0^s K(\eta, -v_n(\eta))d\eta - \int_0^s K(\eta, -v_0(\eta))d\eta\right]\right\} ds$$

$$\leq \left\|\int_0^1 \frac{1}{\Gamma(\alpha)}(1-s)^{\alpha-1}\{|f(s, -v_n(s)) - f(s, -v_0(s))|\right.$$

$$\left. + |\int_0^s K(\eta, -v_n(\eta))d\eta - \int_0^s K(\eta, -v_0(\eta))d\eta|\} \, ds\right\|$$

$$\to 0, \ n \to \infty.$$

By using the same method, we also get $\|T_2 u_n(t) - T_2 u_0(t)\| \to 0$, $n \to \infty$. Which implying that $T : U \to U$ is continuous.

Therefore, T is a completely continuous operator. Since U is a bounded closed convex set. By using the Schauder fixed point theorem (cf [6]), T has a fixed point (u_0, v_0) in U and $T(u_0, v_0) = (u_0, v_0) = (T_1 v_0, T_2 u_0)$.

Let $(u^*, v^*) = -(u_0, v_0)$. Then we have

$$u^* = -u_0(t) = -T_1 v_0$$

$$= \int_0^1 G_\alpha(t,s) \left[f(s, -v_0(s)) + \int_0^s K(\eta, -v_0(\eta)) d\eta \right] ds$$

$$= \int_0^1 G_\alpha(t,s) \left[f(s, v^*(s)) + \int_0^s K(\eta, v^*(\eta)) d\eta \right] ds$$

$$= \int_0^1 G_\alpha(t,s) h_{v^*}(s) ds.$$

$$u^* = -u_0(t) = -T_1 v_0$$

$$= \int_0^1 G_\beta(t,s) \left[g(s, -u_0(s)) + \int_0^s H(\eta, -u_0(\eta)) d\eta \right] ds$$

$$= \int_0^1 G_\beta(t,s) \left[g(s, u^*(s)) + \int_0^s H(\eta, u^*(\eta)) d\eta \right] ds = \int_0^1 G_\beta(t,s) j_{u^*}(s) ds.$$

By **Lemma 2**, it means that (u^*, v^*) is a solution of the problem (1).

Now we turn to show the uniqueness of the solution. Assume that $(u_1(t), v_1(t))$ and $(u_2(t), v_2(t)) \in U$ are two solutions of (1). Then by (2), we have

$$u_1(t) - u_2(t) = \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} \{ [f(s, v_1(s)) - f(s, v_2(s))]$$

$$+ \int_0^s [K(\eta, v_1(\eta)) - K(\eta, v_2(\eta))] d\eta \} ds.$$

Therefore, by (3), for any $0 \le t \le 1$,

$$u_1(t) - u_2(t) \le \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} \left[f_1(|v_1 - v_2|) + \int_0^s f_2(|v_1 - v_2|) d\eta \right] ds$$

$$= \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} f_1(|v_1 - v_2|) ds + \int_0^t \frac{(t-s)^{\alpha-1}}{\Gamma(\alpha)} f_2(|v_1 - v_2|) ds$$

$$\le \frac{2^\alpha}{\Gamma(\alpha)} \int_0^t [f_1(|v_1 - v_2|) + f_2(|v_1 - v_2|)] ds$$

$$\le \frac{2^\alpha}{\Gamma(\alpha)} \int_0^t F_1(|u_1 - u_2| + |v_1 - v_2|) ds.$$

By the similar way,

$$|v_1(t) - v_2(t)| \le \frac{2^\beta}{\Gamma(\beta)} \int_0^t F_2(|u_1 - u_2| + |v_1 - v_2|) ds.$$

We hence have

$$|u_1(t) - u_2(t)| + |v_1(t) - v_2(t)| \leq C \int_0^t (F_1 + F_2)(|u_1 - u_2| + |v_1 - v_2|) ds,$$

where $C = \max\left\{ \frac{2^\alpha}{\Gamma(\alpha)}, \frac{2^\beta}{\Gamma(\beta)} \right\}$.

Assume that

$$G(r) = \int_1^r \frac{1}{(F_1 + F_2)(s)} ds, \ r \in [0, +\infty).$$

Then

$$G(0) = \int_1^0 \frac{1}{(F_1 + F_2)(s)} ds = -\infty, \ G^{-1}(-\infty) = 0.$$

By using the generalized Gronwall inequality [7] we get

$$0 \leq |u_1(t) - u_2(t)| + |v_1(t) - v_2(t)|$$
$$\leq G^{-1}\left[G(0) + \int_0^t \max\left\{ \frac{2^\alpha}{\Gamma(\alpha)}, \frac{2^\beta}{\Gamma(\beta)} \right\} ds \right]$$
$$= G^{-1}(-\infty) = 0.$$

Then

$$u_1(t) = u_2(t), \ v_1(t) = v_2(t).$$

This completes the proof of **Theorem 1**.

References

1. Kilbas A, Srivastava H, Trujillo J (2006) Theory and applications of fractional differential equations. Elsevier B V, New York
2. Feng W, Sun S (2011) Existence of solutions for the system of nonlinear fractional differential equations. J Univ JiNan (Sci and Tech) 25:319–321
3. Xu X, Sun X, Lv W (2011) Existence of positive solution for boundary value problems with nonlinear fractional differential equations. Acta Mathematica Scientia 31A:401–409
4. Yang L, Chen H (2011) Unique positive solution of boundary value problem for fractional differential equations. J Biomathematics 26(1):43–47
5. Yosida K (1999) Functional analysis. Springer, Berlin
6. Istratescu V (2009) Fixed point theory and introduction. D. Reidel publishing company, Dordrecht
7. Lipovan O (2000) A retarded Gronwall-like inequality and its applications. J Math Anal Appl 252:389–401

On the Existence of Extremal Positive Definite Solutions of a Class of Nonlinear Matrix Equation

Bo Wang and Qingchun Li

Abstract The nonlinear matrix equation $X^r + \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i = Q$ is studied. A necessary condition for the existence of positive definite solutions of this equation is derived. Based on the Banach fixed point theorem, a sufficient condition for the existence of a unique positive definite solution of this equation is also derived. Iterative methods for obtaining the extremal (maximal–minimal) positive definite solutions of this equation are proposed.

Keywords Nonlinear matrix equation · Positive definite · Extremal positive solution · Iterative methods

1 Introduction

Consider the nonlinear matrix equation

$$X^r + \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i = Q \, (-1 < \delta_i < 0) \tag{1.1}$$

where A_i are $n \times n$ nonsingular complex matrices, Q is a positive definite matrix, r, m are positive integers.

Several authors [1–6] have studied the existence, the rate of convergence as well as the necessary and sufficient conditions of the existence of positive definite solutions of similar kinds of nonlinear matrix equations. Liu and Gao [7] have studied the properties and sensitivity analysis of the Hermitian positive definite solutions of Eq. (1.1) with the case $m = 1, \delta_i = -t, Q = I$. Duan and Liao [8] and Yueting [9] have studied the matrix equation with the case $m = 1, \delta_i = -t$.

B. Wang · Q. Li (✉)
College of Mathematics, Beihua University, Jilin 132013, People's Republic of China
e-mail: liqingchun01@163.com

This paper is organized as follows: In Sect. 2, we give some notations and lemmas that will be needed to develop this work. We propose a necessary condition for the existence of positive definite solutions of Eq. (1.1), also, we derive a sufficient condition for the existence of a unique positive definite solution of this equation in Sect. 3. In Sect. 4, We propose iterative methods for obtaining extremal maximal- minimal positive definite solutions of Eq. (1.1).

2 Preliminaries

The following properties and the lemmas stated below will be needed for this paper:

1. We use $\lambda_1(B)$ and $\lambda_n(B)$ to denote the maximal and minimal eigenvalues of an $n \times n$ HPD matrix B respectively.
2. For $B = (b_1, b_2, \cdots b_n) = (b_{ij})$ and a matrix C, $B \otimes C = (b_{ij}C)$ is a kronecker product and $vec(B)$ is a vector defined by $vec(B) = (b_1^T, b_2^T, \ldots b_n^T)^T$. We use $\|vec(X)\| = \|X\|_F$, $vec(AXB) = (B^T \otimes A)vec(X)$ where A, X and B are $n \times n$ complex matrices.

Lemma 2.1 [8] *If $A \geq B > 0$ (or $A > B > 0$), then $A^\gamma \geq B^\gamma > 0$ (or $A^\gamma > B^\gamma > 0$) for all $\gamma \in (0, 1]$, and $B^\gamma \geq A^\gamma > 0$ (or $B^\gamma > A^\gamma > 0$) for all $\gamma \in [-1, 0)$.*

3 A Necessary Condition for the Existence of a Positive Definite Solution of $X^r + \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i = Q$

Theorem 3.1 *If Eq. (1.1) has a HPD solution X and A_i are complex nonsingular matrices, Q is a positive definite matrix, then*

$$X \in \left(\frac{1}{m} \sum_{i=1}^{m} \left(A_i Q^{-1} A_i^*\right)^{-\frac{1}{\delta_i}}, Q^{\frac{1}{r}} \right).$$

Proof Let Eq. (1.1) have a HPD solution X. Then

$$X^r = Q - \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i < Q,$$

so $X < Q^{\frac{1}{r}}$. $\square$

Also, $\sum_{i=1}^{m} A_i^* X^{\delta_i} A_i < Q$ implies $A_i^* X^{\delta_i} A_i < Q, \forall i = 1, 2, \ldots m$, hence, $X^{-\delta_i} > A_i Q^{-1} A_i^*$, which leads to $X > \left(A_i Q^{-1} A_i^*\right)^{-\frac{1}{\delta_i}} \forall i = 1, 2, \ldots m$.

Thus,

$$mX > \sum_{i=1}^{m} (A_i Q^{-1} A_i^*)^{-\frac{1}{\delta_i}} \text{ and } X > \frac{1}{m} \sum_{i=1}^{m} (A_i Q^{-1} A_i^*)^{-\frac{1}{\delta_i}},$$

which implies that

$$X \in \left(\frac{1}{m} \sum_{i=1}^{m} (A_i Q^{-1} A_i^*)^{-\frac{1}{\delta_i}}, Q^{\frac{1}{r}} \right)$$

Lemma 3.2 *Let X and Y be positive definite matrices, such that $X \geq aI, Y \geq aI$ for some positive number a. Then*

$$\left\| Y^{\delta} - X^{\delta} \right\| \leq -\delta a^{\delta-1} \|X - Y\|,$$

where $-1 < \delta < 0$.

Proof Put $\gamma = -\delta$, then $0 < \gamma < 1$.

$$Y^{\delta} - X^{\delta} = Y^{-\gamma} - X^{-\gamma} = X^{-\gamma}(X^{\gamma} - Y^{\gamma})Y^{-\gamma},$$
$$\left\| Y^{\delta} - X^{\delta} \right\| \leq \|X^{-\gamma}\| \|X^{\gamma} - Y^{\gamma}\| \|Y^{-\gamma}\|.$$

By Theorem X.3.8. p. 304 [10] we have

$$\left\| Y^{\delta} - X^{\delta} \right\| \leq \gamma a^{\gamma-1} \|X^{-\gamma}\| \|X - Y\| \|Y^{-\gamma}\|.$$

Since, $X \geq aI, Y \geq aI$, then

$$\left\| Y^{\delta} - X^{\delta} \right\| \leq \gamma a^{-\gamma-1} \|X - Y\|.$$

Hence,

$$\left\| Y^{\delta} - X^{\delta} \right\| \leq -\delta a^{\delta-1} \|X - Y\|.$$

Theorem 3.3 *If $\sum_{i=1}^{m} A_i^* X^{\delta_i} A_i \leq (1 - a^r)Q \leq Q - \frac{1}{m^r} \left(\sum_{i=1}^{m} (A_i Q^{-1} A_i^*)^{-\frac{1}{\delta_i}} \right)^r$ for all $\frac{1}{m} \sum_{i=1}^{m} (A_i Q^{-1} A_i^*)^{-\frac{1}{\delta_i}} \leq a Q^{\frac{1}{r}} \leq X \leq Q^{\frac{1}{r}}$, where a is a sufficiently small positive number,*

$$q = m^{r-1} \sum_{i=1}^{m} \left(-\delta_i (a\lambda_n(Q^{\frac{1}{r}}))^{\delta_i-1} \|A_i\|_F^2 \right) \Big/ \left(r \left(\sum_{i=1}^{m} \lambda_n^{-1/\delta_i}(A_i Q^{-1} A_i^*) \right)^{r-1} \right) < 1,$$

then Eq. (1.1) has a unique HPD solution.

Proof Consider the map $f(X) = (Q - \sum_{i=1}^{m} A_i^* Q^{-1} A_i)^{1/\delta_i}$, let $X \in \Omega = \{X | \frac{1}{m} \sum_{i=1}^{m} (A_i Q^{-1} A_i^*)^{-1/\delta_i} \leq a Q^{1/r} \leq X \leq Q^{1/r}\}$.

Obviously, Ω is a convex, closed and bounded set and $f(X)$ is continuous on Ω.

If $\sum_{i=1}^{m} A_i^* X^{\delta_i} A_i \leq (1 - a^r)Q \leq Q - \frac{1}{m^r} (\sum_{i=1}^{m} (A_i Q^{-1} A_i^*)^{-1/\delta_i})^r$ for all $X \in \Omega$, then

$$Q^{\frac{1}{r}} \geq \left(Q - \sum_{i=1}^{m} (A_i^* X^{\delta_i} A_i)^{\frac{1}{r}}\right) \geq (a^r Q)^{\frac{1}{r}} \geq \left[Q - Q + \frac{1}{m^r}\left(\sum_{i=1}^{m} (A_i Q^{-1} A_i^*)^{-\frac{1}{\delta_i}}\right)^r\right]^{\frac{1}{r}},$$

i.e.

$$\frac{1}{m}\sum_{i=1}^{m} (A_i Q^{-1} A_i^*)^{-\frac{1}{\delta_i}} \leq a Q^{\frac{1}{r}} \leq f(X) \leq Q^{\frac{1}{r}}.$$

Hence $f(\Omega) \subseteq \Omega$.

By fixed point theorem, $f(X)$ has a fixed point in $[\frac{1}{m}\sum_{i=1}^{m} (A_i Q^{-1} A_i^*)^{-\frac{1}{\delta_i}}, Q^{\frac{1}{r}}]$.
For arbitrary $X, Y \in \Omega$, we have

$$\sum_{i=1}^{m} A_i^* X^{\delta_i} A_i \leq Q - \frac{1}{m^r}\left(\sum_{i=1}^{m} (A_i Q^{-1} A_i^*)^{-\frac{1}{\delta_i}}\right)^r,$$
$$\sum_{i=1}^{m} A_i^* Y^{\delta_i} A_i \leq Q - \frac{1}{m^r}\left(\sum_{i=1}^{m} (A_i Q^{-1} A_i^*)^{-\frac{1}{\delta_i}}\right)^r$$

i.e.

$$Q - \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i \geq \frac{1}{m^r}\left(\sum_{i=1}^{m} (A_i Q^{-1} A_i^*)^{-\frac{1}{\delta_i}}\right)^r,$$
$$Q - \sum_{i=1}^{m} A_i^* Y^{\delta_i} A_i \geq \frac{1}{m^r}\left(\sum_{i=1}^{m} (A_i Q^{-1} A_i^*)^{-\frac{1}{\delta_i}}\right)^r.$$

Hence,

$$f(X) = (Q - \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i)^{\frac{1}{r}} \geq \frac{1}{m}\sum_{i=1}^{m} (A_i Q^{-1} A_i^*)^{-\frac{1}{\delta_i}} \geq \frac{1}{m}\sum_{i=1}^{m} \lambda_n^{-\frac{1}{\delta_i}}(A_i Q^{-1} A_i^*) I,$$
$$f(Y) = (Q - \sum_{i=1}^{m} A_i^* Y^{\delta_i} A_i)^{\frac{1}{r}} \geq \frac{1}{m}\sum_{i=1}^{m} (A_i Q^{-1} A_i^*)^{-\frac{1}{\delta_i}} \geq \frac{1}{m}\sum_{i=1}^{m} \lambda_n^{-\frac{1}{\delta_i}}(A_i Q^{-1} A_i^*) I,$$

$$\begin{aligned}
\|f(X)^r - f(Y)^r\|_F &= \left\|\sum_{j=0}^{r-1} f(X)^j (f(X) - f(Y)) f(Y)^{r-1-j}\right\|_F \\
&= \left\|\sum_{j=0}^{r-1} vec[f(X)^j (f(X) - f(Y)) f(Y)^{r-1-j}]\right\| \\
&= \left\|\sum_{j=0}^{r-1} (f(Y)^{r-1-j} \otimes f(X)^j) vec(f(X) - f(Y))\right\| \\
&\geq \sum_{i=1}^{m} \frac{1}{m^{r-1}}\left(\sum_{i=1}^{m} \lambda_n^{-\frac{1}{\delta_i}}(A_i Q^{-1} A_i^*)\right)^{r-1} \|vec(f(X) - f(Y))\| \\
&= \frac{r}{m^{r-1}}\left(\sum_{i=1}^{m} \lambda_n^{-\frac{1}{\delta_i}}(A_i Q^{-1} A_i^*)\right)^{r-1} \|f(X) - f(Y)\|_F.
\end{aligned}$$

$$\tag{3.1}$$

According to the map f, we have

$$\begin{aligned}
f(X)^r - f(Y)^r &= \left(Q - \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i\right) - \left(Q - \sum_{i=1}^{m} A_i^* Y^{\delta_i} A_i\right) \\
&= \sum_{i=1}^{m} A_i^* (Y^{\delta_i} - X^{\delta_i}) A_i
\end{aligned} \tag{3.2}$$

Combining (3.1) and (3.2), we have

$$\|f(X) - f(Y)\|_F \leq \frac{m^{r-1}}{r\left(\sum_{i=1}^{m} \lambda_n^{-1/\delta_i}(A_i Q^{-1} A_i^*)\right)^{r-1}} \|f(X)^r - f(Y)^r\|_F$$

$$= \frac{m^{r-1}}{r\left(\sum_{i=1}^{m} \lambda_n^{-1/\delta_i}(A_i Q^{-1} A_i^*)\right)^{r-1}} \left\|\sum_{i=1}^{m} A_i^*(Y^{\delta_i} - X^{\delta_i})A_i\right\|_F,$$

$$\|f(X) - f(Y)\|_F \leq \frac{m^{r-1}}{r\left(\sum_{i=1}^{m} \lambda_n^{-1/\delta_i}(A_i Q^{-1} A_i^*)\right)^{r-1}} \sum_{i=1}^{m} \|A_i\|_F^2 \|Y^{\delta_i} - X^{\delta_i}\|_F.$$

Notice that $X, Y \geq a Q^{\frac{1}{r}} \geq a \lambda_n(Q^{\frac{1}{r}})I$, according to Lemma 3.2, we have

$$\|f(X) - f(Y)\|_F \leq \frac{m^{r-1}}{r\left(\sum_{i=1}^{m} \lambda_n^{-1/\delta_i}(A_i Q^{-1} A_i^*)\right)^{r-1}} \sum_{i=1}^{m} -\delta_i(a \lambda_n(Q^{\frac{1}{r}}))^{\delta_i - 1} \|A_i\|_F^2 \|X - Y\|_F.$$

Since $q < 1$, $f(X)$ is a contraction map in Ω. By Banach's fixed point theorem, the map $f(X)$ has a unique fixed point in Ω and this shows that Eq. (1.1) has a unique HPD solution in $\left[\frac{1}{m}\sum_{i=1}^{m}(A_i Q^{-1} A_i^*)^{-\frac{1}{\delta_i}}, Q^{\frac{1}{r}}\right]$.

4 Iterative Algorithms for Obtaining Extremal Positive Definite Solutions of $X^r + \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i = Q$

In this section, we consider iterative algorithms for obtaining extremal positive definite solutions (if they exist) to Eq. (1.1).

Algorithm 4.1 Consider the iterative algorithm

$$X_0 = \alpha Q^{\frac{1}{r}}, \alpha > 1$$

$$X_{k+1} = \left(Q - \sum_{i=1}^{m} A_i^* X_k^{\delta_i} A_i\right)^{\frac{1}{r}}, k = 0, 1, \ldots \tag{4.1}$$

Theorem 4.2 *If $A_i, i = 1, 2, \ldots, m$ A_i are complex nonsingular matrices, Q is a positive definite matrix and there is a real number $\alpha(\alpha > 1)$, the following condition is satisfied $\sum_{i=1}^{m} \alpha^{\delta_i} A_i^* Q^{\frac{\delta_i}{r}} A_i > (1 - \alpha^r)Q$, and if Eq. (1.1) has a positive definite solution X, then the sequence of positive definite matrices $\{X_k\}$ derived from Algorithm 4.1 is monotonic decreasing and bounded from below and hence converges to X_L (the maximal solution).*

Proof We consider the sequence of matrices generated by Algorithm 4.1

For $k = 0$ we have

$$X_1 = \left(Q - \sum_{i=1}^{m} A_i^* X_0^{\delta_i} A_i\right)^{\frac{1}{r}} = \left(Q - \sum_{i=1}^{m} \alpha^{\delta_i} A_i^* Q^{\frac{\delta_i}{r}} A_i\right)^{\frac{1}{r}}.$$

Using the condition of the theorem, we obtain $Q - \sum_{i=1}^{m} \alpha^{\delta_i} A_i^* Q^{\frac{\delta_i}{r}} A_i < \alpha^r Q$, so we get $(Q - \sum_{i=1}^{m} \alpha^{\delta_i} A_i^* Q^{\frac{\delta_i}{r}} A_i)^{\frac{1}{r}} < \alpha Q^{\frac{1}{r}}$, i.e. $X_1 < X_0$.

For $k = 1$,

$$X_2 = \left(Q - \sum_{i=1}^{m} A_i^* X_1^{\delta_i} A_i\right)^{\frac{1}{r}} = \left(Q - \sum_{i=1}^{m} A_i^* \left(Q - \sum_{i=1}^{m} \alpha^{\delta_i} A_i^* Q^{\frac{\delta_i}{r}} A_i\right)^{\frac{\delta_i}{r}} A_i\right)^{\frac{1}{r}}.$$

Since $Q - \sum_{i=1}^{m} \alpha^{\delta_i} A_i^* Q^{\frac{\delta_i}{r}} A_i < \alpha^r Q$, then $(Q - \sum_{i=1}^{m} \alpha^{\delta_i} A_i^* Q^{\frac{\delta_i}{r}} A_i)^{\frac{\delta_i}{r}} > \alpha^{\delta_i} Q^{\frac{\delta_i}{r}}$ and $A_i^*(Q - \sum_{i=1}^{m} \alpha^{\delta_i} A_i^* Q^{\frac{\delta_i}{r}} A_i)^{\frac{\delta_i}{r}} A_i > \alpha^{\delta_i} A_i^* Q^{\frac{\delta_i}{r}} A_i, \forall i = 1, 2, 3 \ldots$, therefore, $(Q - \sum_{i=1}^{m} A_i^*(Q - \sum_{i=1}^{m} \alpha^{\delta_i} A_i^* Q^{\frac{\delta_i}{r}} A_i)^{\frac{\delta_i}{r}} A_i)^{\frac{1}{r}} < (Q - \sum_{i=1}^{m} \alpha^{\delta_i} A_i^* Q^{\frac{\delta_i}{r}} A_i)^{\frac{1}{r}} < \alpha Q^{\frac{1}{r}}$, so $X_2 < X_0$.

Since $X_1 < X_0$, then $X_1^{\delta_i} > X_0^{\delta_i}$ and $\sum_{i=1}^{m} A_i^* X_1^{\delta_i} A_i > \sum_{i=1}^{m} A_i^* X_0^{\delta_i} A_i$.

Therefore, $(Q - \sum_{i=1}^{m} A_i^* X_1^{\delta_i} A_i)^{\frac{1}{r}} < (Q - \sum_{i=1}^{m} A_i^* X_0^{\delta_i} A_i)^{\frac{1}{r}}$, so $X_2 < X_1$, and hence $X_0 > X_1 > X_2$.

We receive by analogy $X_1 > X_3$ and $X_2 > X_3$. Consequently, $X_0 > X_1 > X_2 > X_3$.

Analogously we can prove:

$$X_0 > X_1 > X_2 > X_3 > X_4 \ldots$$

Thus, $\{X_k\}$ is a monotonic decreasing sequence. Next, we show that $\{X_k\}$ is bounded from below by some positive definite solution $X(X_k > X)$ of Eq. (1.1). This can be proved using an induction as follows:

For $k = 0$,

$$X_0 - X = \alpha Q^{\frac{1}{r}} - (Q - \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i)^{\frac{1}{r}}.$$

Since $Q - \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i < Q$, then $(Q - \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i)^{\frac{1}{r}} < Q^{\frac{1}{r}}$ and

$$\alpha Q^{\frac{1}{r}} - \left(Q - \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i\right)^{\frac{1}{r}} > \alpha Q^{\frac{1}{r}} - Q^{\frac{1}{r}} = (\alpha - 1) Q^{\frac{1}{r}} > 0.$$

That is, $X_0 - X > 0$, i.e. $X_0 > X$.

Assume that the inequality $X_k > X$ holds at $k = t$ i.e. $X_t > X$.

Now, For $k = t + 1$ we have

$$X_{t+1} - X = \left(Q - \sum_{i=1}^{m} A_i^* X_t^{\delta_i} A_i\right)^{\frac{1}{r}} - \left(Q - \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i\right)^{\frac{1}{r}}.$$

Since $X_t > X$, then $X_t^{\delta_i} < X^{\delta_i}$. Therefore, $\sum_{i=1}^{m} A_i^* X_t^{\delta_i} A_i < \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i$ and $(Q - \sum_{i=1}^{m} A_i^* X_t^{\delta_i} A_i)^{\frac{1}{r}} > (Q - \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i)^{\frac{1}{r}}$

Hence $X_{t+1} - X > 0$, i.e. $X_{t+1} > X$ as required.

Therefore, $\{X_k\}$ is a monotonic decreasing sequence and is bounded from below by some positive definite solution X, then it converges to a positive definite solution matrix X. Taking the limit of (4.1) shows that the sequence converges in the set of solutions and hence X is a solution. But, since X is an arbitrarily chosen positive definite solution of Eq. (1.1) so we get $X_k > X_L, \forall k \in N$. Hence X_k converges to X_L.

Algorithm 4.3 Consider the iterative algorithm

$$X_0 = \alpha Q^{\frac{1}{r}}, 0 < \alpha < 1$$

$$X_{k+1} = (Q - \sum_{i=1}^{m} A_i^* X_k^{\delta_i} A_i)^{\frac{1}{r}}, k = 0, 1, \ldots$$

Similar the proof of the Theorem 4.2, we can obtain the following theorem.

Theorem 4.4 *If A_i, $i = 1, 2, \ldots, m$ A_i are complex nonsingular matrices, Q is a positive definite matrix and there is a real number $\alpha(0 < \alpha < 1)$, and the following condition is satisfied*

$$\sum_{i=1}^{m} \alpha^{\delta_i} A_i^* Q^{\frac{\delta_i}{r}} A_i < (1 - \alpha^r)Q,$$

and if Eq. (1.1) has a positive definite solution $\tilde{X}$, then the sequence of positive definite matrices $\{X_k\}$ derived from Algorithm 4.3 is monotonic increasing and bounded from below and hence converges to X_s (the minimal solution).

References

1. Duan X, Liao A, Tang B (2008) On the nonlinear matrix equation $X - \sum_{i=1}^{m} A_i^* X^{\delta_i} A_i = Q$. Linear Algebra Appl 429:110–121
2. El-Sayed SM, Ramadan MA (2001) On the existence of a positive definite solution of the matrix equation $X - A^* \sqrt[2^m]{X^{-1}} A = I$. Int J Comput Math 76:331–338
3. Peng ZY, El-sayed SM, Zhang XL (2007) Iterative method for the extremal positive definite solution of the matrixequation $X + A^* X^{-\alpha} A = Q$. Appl Math Comput 200:520–527
4. Mohamed A (2005) Ramadan: on the existence of extremal positive definite solutions of a kind of matrix. Int J Nonlinear Sci Numer Simul 6:115–126
5. Ramadan MA (2005) Necessary and sufficient conditions to the existence of positive definite solution of the matrix equation $X + A^T X^{-2} A = I$. Int J Comput Math 82:865–870
6. Zhan X (1996) Computing the extremal positive definite solutions of a matrix equations. SIAM J Sci Comput 17:1167–1174
7. Liu XG, Gao H (2003) On the positive definite solutions of a matrix equations $X^s \pm A^T X^{-t} A = I_n$. Linear Algebra Appl 368:83–97
8. Duan X, Liao A (2008) On the existence of hermitian positive definite solutions of the matrix equation $X^s + A^* X^{-t} A = Q$. Linear Algebra Appl 429:673–687

9. Yueting Y (2007) The iterative method for solving nonlinear matrix equation $X^s + A^*X^{-t}A = Q$. Appl Math Comput 188:46–53
10. Bhatia R (1997) Matrix analysis, vol 169. Springer, Berlin

An Iterative Algorithm
for the Generalized Center Symmetric
Solutions of a Class of Linear Matrix
Equation and Its Optimal Approximation

Jie Liu and Qingchun Li

Abstract For any symmetric orthogonal matrix P, i.e., $P^{\mathrm{T}} = P$, $P^{\mathrm{T}}P = I$, the matrix X is said to be a generalized centrosymmetric matrix if $PXP = X$ for any matrix X. The conjugate gradient iteration algorithm is presented to find the generalized centrosymmetric solution and its optimal approximation of the constraint matrix equation $AXB + CXD = F$. By this method, the solvability of the equation can be determined automatically. If the matrix equation $AXB + CXD = F$ is consistent, then its generalized centrosymmetric solution can be obtained within finite iteration steps in the absence of round off errors for any initial symmetric matrix X_1, and generalized centrosymmetric solution with the least norm can be derived by choosing a proper initial matrix. In addition, the optimal approximation solution for a given matrix of the matrix equation $AXB + CXD = F$ can be obtained by choosing the generalized centrosymmetric solution with the least norm of a new matrix equation $A\tilde{X}B + C\tilde{X}D = \tilde{F}$.

1 Introduction

$R^{m \times n}$ is the set of $m \times n$ real matrices, A^{T} is the transpose of a matrix A, I_n is the unit matrix with order n, $GCSR_P^{n \times n}$ is the set of the generalized centrosymmetric matrix. We define the inner product of two matrices $A, B \in R^{m \times n}$ as $\langle A, B \rangle = tr(B^{\mathrm{T}}A)$, Then the matrix norm of A induced by this inner product is Frobenius norm and is denoted by $\|A\|$, i.e. $\|A\| = \sqrt{\langle A, A \rangle} = tr(A^{\mathrm{T}}A)^{\frac{1}{2}}$. Define vector

J. Liu · Q. Li (✉)
College of Mathematics, Beihua University, Jilin 132013, People's Republic of China
e-mail: liqingchun01@163.com

$vecA = (a_1^T, a_2^T, \cdots, a_n^T)^T$, where $A = (a_1, a_2, \cdots, a_n)$, $A \in R^{m \times n}$, $a_i \in R^m$, $i = 1, 2, \cdots n$. $A \otimes B$ is the Kronecker product of two matrices A and B.

Definition 1 For any given matrix $P \in SOR^{n \times n}$, i.e. $P^T = P$, $P^T P = I$, the matrix $A \in R^{n \times n}$ is said to be a generalized centrosymmetric matrix if $PAP = A$. $GCSR_P^{n \times n}$, is the set of the generalized centrosymmetric matrices.

In this paper, we consider two problems as follows:

Problem 1 Given $A, C \in R^{m \times n}$, $B, D \in R^{n \times p}$, $F \in R^{m \times p}$, $P \in SOR^{n \times n}$, find $X \in GCSR_P^{n \times n}$, such that

$$AXB + CXD = F \tag{1}$$

Problem 2 If the problem 1 is consistent, let S_E denote the set of its solutions, for given $X_0 \in R^{n \times n}$, find $\hat{X} \in S_E$, such that

$$\left\| \hat{X} - X_0 \right\| = \min_{X \in S_E} \| X - X_0 \| \tag{2}$$

For the problem of solving matrix equations, Scholars at home and abroad has obtained a lot of important results [1–3]. For the matrix equation $AXB = C$, Peng [4] discussed its symmetric(skew-symmetric) solutions, centrosymmetric (centrally skew-symmetric) solutions, reflexive(anti-reflexive) matrix solutions, bisymmetric(symmetric and anti-persymmetric) solutions and its optimal approximation; Liang [5] discussed its generalized (anti-symmetric) per-symmetric solutions, generalized (center anti-symmetric) centrosymmetric solutions, generalized bisymmetric(symmetric and anti-persymmetric) solutions and its optimal approximation; Wu [6] discussed its generalized reflexive(anti-reflexive) solutions. In addition, Liang and Liu [7] discussed the iterative algorithms for the minimum-norm solution and the least squares solution of the linear matrix equations $A_1 X B_1 + C_1 X^T D_1 = M_1$, $A_2 X B_2 + C_2 X^T D_2 = M_2$; Ding et al. [8] discussed the iterative solutions to matrix equations of the form $A_i X B_i = F_i$. For the matrix equation $AXB + CXD = F$, Dajin et al. [9] discussed its centrosymmetric solutions and its optimal approximation; Hailin [10] discussed the iterative algorithm for obtaining its symmetric solutions. In this paper, the conjugate gradient iteration algorithm is presented to find the generalized centrosymmetric solution and its optimal approximation of the constraint matrix equation $AXB + CXD = F$.

2 The Solutions of Problem 1

Algorithm 1
(1) Input matrices $A, C \in R^{m \times n}$, $B, D \in R^{n \times p}$, $F \in R^{m \times p}$, $P \in SOR^{n \times n}$, and an initial matrix $X_1 \in GCSR_P^{n \times n}$;

(2) Calculate

$$R_1 = F - (AX_1B + CX_1D), \quad P_1 = A^T R_1 B^T + C^T R_1 D^T,$$

$$Q_1 = \frac{1}{2}(P_1 + PP_1P), \quad k := 1;$$

(3) Calculate

$$X_{k+1} = X_k + \frac{\|R_k\|^2}{\|Q_k\|^2} Q_k;$$

(4) Calculate

$$R_{k+1} = F - (AX_{k+1}B + CX_{k+1}D), \quad P_{k+1} = A^T R_{k+1} B^T + C^T R_{k+1} D^T,$$

$$Q_{k+1} = \frac{1}{2}(P_{k+1} + PP_{k+1}P) - \frac{tr(P_{k+1}^T Q_k)}{\|Q_k\|^2} Q_k;$$

If $R_{k+1} = 0$ or $R_{k+1} \neq 0$, $Q_{k+1} = 0$ then stop; else, $k := k + 1$, go to Step (3).

Lemma 1 *The sequences* $\{R_i\}, \{Q_i\}, \{P_j\}, i, j = 1, 2, \ldots$ *generated by Algorithm 1 satisfy the following relations*

$$tr(R_{i+1}^T R_j) = tr(R_i^T R_j) - \frac{\|R_i\|^2}{\|Q_i\|^2} tr(Q_i^T P_j). \tag{3}$$

Proof

$$tr(R_{i+1}^T R_j) = tr[(F - (AX_{i+1}B + CX_{i+1}D))^T R_j]$$

$$= tr\left[\left(F - A\left(X_i + \frac{\|R_i\|^2}{\|Q_i\|^2} Q_i\right)B - C\left(X_i + \frac{\|R_i\|^2}{\|Q_i\|^2} Q_i\right)D\right)^T R_j\right]$$

$$= tr\left[\left(R_i^T - \frac{\|R_i\|^2}{\|Q_i\|^2}(AQ_iB + CQ_iD)^T\right)R_j\right]$$

$$= tr(R_i^T R_j) - \frac{\|R_i\|^2}{\|Q_i\|^2} tr[(AQ_iB + CQ_iD)^T R_j]$$

$$= tr(R_i^T R_j) - \frac{\|R_i\|^2}{\|Q_i\|^2} tr[Q_i^T (A^T R_j B^T + C^T R_j D^T)]$$

$$= tr(R_i^T R_j) - \frac{\|R_i\|^2}{\|Q_i\|^2} tr[Q_i^T P_j]$$

Using mathematical induction and Lemma 1, we can establish the following two lemmas.

Lemma 2 *The sequences $\{R_i\}, \{Q_i\}$ generated by Algorithm* 1 *satisfy the following relations*

$$tr(R_i^{\mathrm{T}}R_j) = 0,\ tr(Q_i^{\mathrm{T}}Q_j) = 0,\quad i, j = 1, 2, \ldots, k,\ i \neq j. \tag{4}$$

Lemma 3 *Suppose that $\bar{X}$ is a solution of Problem* 1, *then*

$$tr[(\bar{X} - X_k)^{\mathrm{T}}Q_k] = \|R_k\|^2,\ k = 1, 2, \ldots. \tag{5}$$

Theorem 1 *Suppose that Problem* 1 *is consistent, for any initial matrix $X_1 \in GCSR_P^{n \times n}$, the generalized centrosymmetric solution of* (1) *can be obtained within finite iteration steps.*

Proof Suppose that $R_i \neq 0,\ i = 1, 2, \ldots, mp$ by Lemma 3 we have $Q_i \neq 0,\ i = 1, 2, \ldots, mp$, therefore X_{mp+1}, R_{mp+1} can be computed by Algorithm 1. By Lemma 2, $tr(R_{mp+1}^{\mathrm{T}}R_i) = 0$, and $tr(R_i^{\mathrm{T}}R_j) = 0,\quad i, j = 1, 2, \ldots, mp, i \neq j$, then $R_1, R_2, \ldots, R_{mp}$ are orthogonal basis of $R^{m \times p}$, and then $R_{mp+1} = 0$, therefore X_{mp+1} is a solution of Problem 1.

If Problem 1 is consistent, we can get a solution after at most $t_0 (t_0 = \min(mp,\ n^2))$ steps. Suppose that $R_i \neq 0\ i = 1, 2, \ldots, n^2$, then $Q_i \neq 0,\ i = 1, 2, \ldots, n^2$. We can get X_{n^2+1}, R_{n^2+1} by Algorithm 1. Similarly, we can obtain $Q_{n^2+1} = 0$, and X_{n^2+1} is a solution of Problem 1.

Corollary 1 *The sufficient and necessary condition that makes Problem 1 inconsistent is that it exists a positive integer k, such that $R_k \neq 0$ but $Q_k = 0$.*

Proof If Problem 1 is inconsistent, for any positive integer i, $R_i \neq 0$. Suppose that $Q_i \neq 0$ for any positive integer i, then from the proof of Theorem 1, we have that Problem 1 is consistent, which is contradicted with the inconsistency of Problem 1. So there exists a positive integer k, such that $R_k \neq 0$ but $Q_k = 0$.

Conversely, suppose that there exists a positive integer k, such that $R_k \neq 0$ but $Q_k = 0$, then Problem 1 is inconsistent by Lemma 3.

Lemma 4 *Constraint matrix equation* (1) *is consistent if and only if the matrix equations*

$$\begin{cases} AXB + CXD = F \\ APXPB + CPXPD = F \end{cases} \tag{6}$$

is consistent.

Proof Suppose that the matrix equation (1) is consistent, and it has a solution $X_0 \in GCSR_P^{n \times n}$, then $X_0 = PX_0P$, $AX_0B + CX_0D = F$ and $APX_0PB + CPX_0PD = F$. Obviously, X_0 is a solution of the matrix equation (6). On the contrary, suppose that the matrix equation (6) is consistent, which has a solution $\bar{X} \in R^{n \times n}$, and $A\bar{X}B + C\bar{X}D = F$, $AP\bar{X}PB + CP\bar{X}PD = F$. Let $\hat{X} = \frac{\bar{X} + P\bar{X}P}{2}$. Then $\hat{X} \in GCSR_P^{n \times n}$ and

$$A\hat{X}B + C\hat{X}D = A\frac{\bar{X} + P\bar{X}P}{2}B + C\frac{\bar{X} + P\bar{X}P}{2}D$$
$$= \frac{A\bar{X}B + C\bar{X}D + (AP\bar{X}PB + CP\bar{X}PD)}{2}$$
$$= \frac{F}{2} + \frac{F}{2}$$
$$= F.$$

Therefore, $\hat{X}$ is a solution of the matrix equation (1), and the matrix equation (1) is consistent.

In conclusion, Lemma 4 has been proved.

Lemma 5 *Peng* [4]. *Assume that $y_0 \in R(M^{\mathrm{T}})$ is a solution of the consistent linear matrix equation $My = b$, then y_0 is the unique least-norm constraint solution.*

Theorem 2 *Suppose that the equation* (1) *is consistent, choosing the initial matrix*

$$X_1 = A^{\mathrm{T}}HB^T + C^{\mathrm{T}}HD^T + PA^{\mathrm{T}}HB^TP + PC^{\mathrm{T}}HD^TP(H \in R^{m \times p}),$$

specially, choosing $X_1 = 0 \in R^{n \times n}$, then the unique least-norm generalized centrosymmetric solution can be obtained within finite iterative steps by Algorithm 1.

Proof From the iterative algorithm and Lemma 1, we choose

$$X_1 = A^{\mathrm{T}}HB^T + C^{\mathrm{T}}HD^T + PA^{\mathrm{T}}HB^TP + PC^{\mathrm{T}}HD^TP(H \in R^{m \times p}),$$

then the solution X^* of Problem 1 is obtained by Algorithm 1, and X^* has the following form

$X_1 = A^{\mathrm{T}}YB^T + C^{\mathrm{T}}YD^T + PA^{\mathrm{T}}YB^TP + PC^{\mathrm{T}}YD^TP.$

From Lemma 4, the generalized centrosymmetric solution in the matrix equation (1) is the solution of the matrix equation (6). So in order to prove that X^* is the least-norm constraint solution of problem 1, we can just prove that X^* is the least-norm constraint solution of the matrix equation (6).

Let $vec(X^*) = x^*, vec(X) = x, vec(Y) = y, vec(F) = f$. The matrix equation(6) is equivalent to the following equations

$$\begin{pmatrix} B^{\mathrm{T}} \otimes A + D^{\mathrm{T}} \otimes C \\ (B^{\mathrm{T}}P) \otimes (AP) + (D^{\mathrm{T}}P \otimes (CP)) \end{pmatrix} x = \begin{pmatrix} f \\ f \end{pmatrix}. \tag{7}$$

Moreover

$$x^* = vec(x^*) = vec(A^{\mathrm{T}}YB^T + C^{\mathrm{T}}YD^T + PA^{\mathrm{T}}YB^TP + PC^{\mathrm{T}}YD^TP)$$
$$= (B \otimes A^{\mathrm{T}} + D \otimes C^{\mathrm{T}})y + [(PB) \otimes (PA^{\mathrm{T}}) + (PD) \otimes (PC^{\mathrm{T}})]y$$
$$= (B \otimes A^{\mathrm{T}} + D \otimes C^{\mathrm{T}}, (PB) \otimes (PA^{\mathrm{T}}) + (PD) \otimes (PC^{\mathrm{T}})) \begin{pmatrix} y \\ y \end{pmatrix}$$

$$= \begin{bmatrix} B^T \otimes A + D^T \otimes C \\ (B^T P) \otimes (AP) + (D^T P) \otimes (CP) \end{bmatrix}^T \begin{pmatrix} y \\ y \end{pmatrix}$$

$$\in R\left(\begin{bmatrix} B^T \otimes A + D^T \otimes C \\ (B^T P) \otimes (AP) + (D^T P) \otimes (CP) \end{bmatrix} \right).$$

From Lemma 5, X^* is the unique least-norm constraint solution of the matrix equation (7), and the operator *vec* is isomorphic, then X^* is the unique least-norm constraint solution of Problem 1.

3 The Solutions of Problem 2

Suppose that the solution set of Problem 1 is non-empty. For any matrix $X \in GCSR_P^{n \times n}$, assumed that $X_0 \in R^{n \times n}$ is given in Problem 2, then

$$\|X - X_0\|^2 = \left\| X - \left(\frac{X_0 + PX_0P}{2} + \frac{X_0 - PX_0P}{2} \right) \right\|^2$$

$$= \left\| \left(X - \frac{X_0 + PX_0P}{2} \right) - \frac{X_0 - PX_0P}{2} \right\|^2$$

$$= \left\| X - \frac{X_0 + PX_0P}{2} \right\|^2 + \left\| \frac{X_0 - PX_0P}{2} \right\|^2,$$

in order to find $\hat{X} \in S_E$ in Problem 2, such that $\|\hat{X} - X_0\| = \min_{x \in S_E} \|X - X_0\|$, which is equivalent to find $\hat{X} \in S_E$, such that

$$\left\| \hat{X} - \frac{X_0 + PX_0P}{2} \right\| = \min_{x \in S_E} \left\| X - \frac{X_0 + PX_0P}{2} \right\|.$$

Notice that Problem 1 is consistent, then for $X_0 \in R^{n \times n}$, there is a matrix $X \in S_E$, such that $AXB + CXD = F$, then

$$A(X - X_0)B + C(X - X_0)D = F - AX_0B - CX_0D.$$

Let $\tilde{X} = X - X_0$ $\tilde{F} = F - AX_0B - CX_0D$. Then the matrix equation (1) is equivalent to

$$A\tilde{X}B + C\tilde{X}D = \tilde{F}. \tag{8}$$

Obviously, solving Problem 2 is translated into solving the least-norm constraint solution $\tilde{X}^*$ of matrix equation (8), according to Algorithm 1, choose the initial matrix $\tilde{X}_1 = A^T HB^T + C^T HD^T + PA^T HB^T P + PC^T HD^T P (H \in R^{m \times p})$, specially, choose $\tilde{X}_1 = 0 \in R^{n \times n}$, the unique least-norm generalized centrosymmetric solution of the matrix equation (8) can be obtained, and then $\hat{X} = \tilde{X}^* + X_0$ is the solution of Problem 2.

References

1. Baksalary JK, Kala R (1980) The matrix equation. Linear Algebra Appl 37:141–147
2. Higham NJ (1988) Computing a nearest symmetric positive semidefinite matrix. Linear Algebra Appl 103:103–118
3. Shim SY, Chen Y (2003) Least squares solution of matrix equation. SIAM J Matrix Anal Appl 24:802–808
4. Peng Y (2004) The iterative method for the solutions and the optimal approximation of the constrained matrix equation. Hunan University, Hunan
5. Liang M (2007) Iterative methods for several constrained matrix equation problems and associated optimal approximation. Lanzhou University, Gansu
6. Wu W (2011) The study of iterative algorithms for solving linear matrix equations. Nanchang University, Jiangxi
7. Liang K, Liu J (2011) The iterative algorithms for the minimum-norm solution and the least squares solution of the linear matrix equations $A_1XB_1 + C_1X^{\mathrm{T}}D_1 = M_1$, $A_2XB_2 + C_2X^{\mathrm{T}}D_2 = M_2$ Appl Math Comput 33:3166–3175
8. Ding J, Liu Y, Ding F (2010) Iterative solutions to matrix equations of the form $A_iXB_i = F_i$ Comput Math Appl 59:3500–3507
9. Dajin L, Hailin Z, Dongjin Y (2008) An iterative algorithm for the centrosymmetric solutions and optimal approximation of $AXB + CXD = F$ J Yangzhou Univ 11:9–13
10. Hailin Z (2010) An iterative method for symmetric solutions of the matrix equation $AXB + CYD = F$ Math Numer Sinica 32:413–422

Multi-Objective Genetic Algorithm with Complex Constraints Based on Colony Classify

Li-li Zhang, Feng Xu and Juan Hu

Abstract The paper presents a constraint-handling approach for multi-objective optimization. The general idea is shown as follow: Firstly, the population was classified into two groups: feasible population and infeasible population. Secondly, feasible population was classified into Pareto population and un-Pareto population. Thirdly, the Pareto population was defied with k-average classify approach into colony Pareto population and in-colony Pareto population. Last, R-fitness was given to each population. Simulation results show that the algorithm not only improves the rate of convergence but also can find feasible Pareto solutions distribute abroad and even.

1 Foreword

Multi-objective optimization problem with complex constraints in real life closely, important for many scholars. Genetic algorithms because the population strategy, what can find multiple Pareto optimal solution,is an effective way to solve the multi-objective optimization problem.In the history the quite classical algorithm includes: Schaffer [1] proposed VEGA, Horn and Nafpliotis [2] proposed NPGA. Fonseca and Fleming [3] proposed FFGA, Srinivas and Dab [4] proposed NSGA,

L. Zhang (✉) · F. Xu
Mathematics Department, School of Science, Anhui University of Science
and Technology, Huainan 232001 Anhui, China
e-mail: xiamilao@126.com

J. Hu
The Foundation Department of Huainan Vocational Technical College,
Huainan 232001, China

Zitzler and Thiele [5] proposed SPEA. The key to solve the problem of multi-objective is to properly handle the constrains. Currently, more evolutionary algorithm for unconstrained optimization problems in the literature, than for constrained optimization problems. In Ref. [6], the current constraint handling methods based on evolutionary algorithms more detailed overview. This paper presents a multi-objective genetic algorithm based group classification with complex constraints.

2 Multi-Objective Optimization Problem Overview

Definition 1 (*Constrained Optimization Problem, COP*) Wang and Liu [7] the constraints multi-objective optimization question to be possible to describe for the following mathematical model:

$$
\begin{aligned}
&\min f(x) = [f_1(x), f_2(x), \cdots, f_p(x)]^T \\
&Subject\,to \\
&\qquad h_i(x) = 0, \ i = 1, \cdots, k \\
&\qquad g_j(x) \leq 0, \ j = 1, \cdots, m
\end{aligned}
\tag{1}
$$

In the formula, $f(x)$ is the goal vector, $h_i(x), g_j(x)$ is called the constraints, $x = (x_1, x_2, \cdots, x_n) \in R^n$ is called the n-dimensional decision vector. Will satisfy all constraints the solution space S to be called (1) the formula feasible territory, min said vector minimization, namely vector goal $f(x) = [f_1(x), f_2(x), \cdots, f_p(x)]^T$ under some kind of constraints each subtarget function minimizing as far as possible. In particular, when p = 1, type (1) into a single objective optimization problem, when $p > 1$, the type (1) was the multi-objective optimization question, when $h_i(x), g_j(x)$ for linearity or nonlinear function, the type (1) to have the complex constraints multi-objective optimization question, this article main discussion has the complex constraints multi-objective optimization question.

Because regarding the equality constraint $h_i(x) = 0$ may through the admissable error (also say that tolerance) $\varepsilon > 0$ transforms it two inequality constraints.

$$
\begin{cases}
h_i(x) - \varepsilon \leq 0 \\
-h_i(x) - \varepsilon \leq 0
\end{cases}
\tag{2}
$$

Therefore may only consider the inequality constraint optimization question.

In the COP optimal process, various goals often is conflicts mutually, one can satisfy all constraints, and the ability to turn all of the objective functions to achieve global optimal solution, it may not exist. So the definition of COP solution, Pareto solution is a key concept.

Definition 2 (*Pareto optimal solution*) Wang and Cao [8] the solution $X^* = \left(x_1^*, x_2^*, \cdots, x_n^*\right) \in S$ is the COP Pareto optimal solution, and when $y \in S$ does not only exist, $\vec{v} = f(y)$ causes to surpass $\vec{u} = f(X^*)$, namely $\forall i \in \{1, \cdots, p\}$, $u_i \leq v_i$ satisfied, and $\exists i \in \{1, \cdots, p\}$ s.t $u_i \prec v_i$.

3 Complex Multi-Objective Genetic Algorithm Which Classifies Based on the Community

3.1 Divide into the Population Feasible and Infeasible Two Kinds

In the genetic algorithm, when the population is large, vector dimension is larger, Vector comparison is troublesome and time consuming. So we can use the vector norm to measure the length of a vector, the vector is converted to vector length comparison comparison. This method is not only effective, but also can greatly improve the efficiency of the algorithm [9].

records the constraints is

$$G_i(x) = \begin{cases} h_i(x) - \varepsilon \leq 0 \\ -h_i(x) - \varepsilon \leq 0 & i = 1, 2, \cdots, N \\ g_i(x) \leq 0 \end{cases}$$

Making $c_i(x) = \max(G_i(x), 0) \quad i = 1, 2, \cdots, N$

Results in the vector

$$C(x) = (c_1(x), c_2(x), \cdots, c_N(x))^T$$

Making

$$D_i(x) = \left(\sum_{i=1}^{M} [c_i(x)]^p \right)^{\frac{1}{p}} \quad (p > 0)$$

From this formula can be seen, $D_i(x)$ actually reflect each individual distance feasible space length: $D_i(x) = 0$, then the individual is located in the feasible space, is a feasible solution; $D_i(x) = D_j(x)$, said the two individuals with the same length distance feasible space; $0 \prec D_i(x) \prec D_j(x)$, said that although $D_i(x)$ and $D_j(x)$ as infeasible solution, but with j compared to i individual, individual distance feasible space is relatively close, namely individual j is superior to individual i. Here $D(x)$ is called D- fitness, using D- fitness hand can realize the search space to the space of feasible mapping, on the other hand the entire colony is divided into viable groups and not feasible groups, more important is to be able
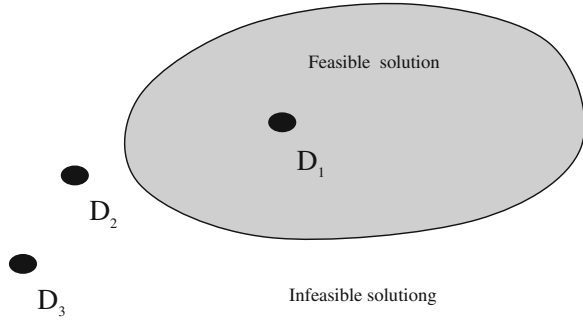
to adapt the D-fitness as the genetic algorithm fitness, selection, crossover and mutation to generate a large number of feasible individual [10] (Fig. 1).

Such complex constrained multi-objective optimization genetic algorithm can be divided into infeasible solutions and feasible solution, which infeasible individual collections formed infeasible groups, feasible individual collections constitute feasible groups. For some constraint conditions many optimization problems, as in previous generations may not produce or generate a feasible solution. If the feasible groups Pareto operations and cluster analysis operation, one side will reduce the efficiency of the algorithm, on the other hand, also do not have much practical significance. This algorithm first step is to generate a large number of feasible solution, then we use the D-fitness as algorithm fitness, through the selection, crossover and mutation, as in previous generations generates as much feasible solution for individual, behind the Pareto operation and lay the foundation for.

3.2 Viable Groups Divided into Viable Non Pareto Group and Pareto Group Feasible

By Pareto optimal solution definition, if a multi-objective optimization problem optimal solution, then the solution must be the Pareto optimal solution, so solving the multi-objective optimization problem is to find out the all Pareto optimal solution. Since the genetic algorithm to the entire population of evolutionary operation, it focuses on the individual set, and the multi objective optimization problem of Pareto optimal solution in general is also a collection, thus the use of genetic algorithm to solve the multi-objective optimization problem of Pareto optimal solution set is a kind of effective means. In this algorithm, the number of individuals when feasible enough, can carries on the Pareto operation, find all the Pareto optimal solution, which will be feasible groups into viable non Pareto group and Pareto group feasible.

3.3 The Feasible Pareto Colony is Divided into Pareto Group and Pareto Clustering Optimal Population

3.3.1 K- Means Clustering

Because the genetic algorithm on the same individual or similar individual number does not limit, so in practical Pareto groups may appear a large number of identical or similar to the individual, especially in the later stages of Pareto algorithm feasible individual has very much, so it is necessary to restrict their number, so as to be capable of generating more abundant species. Different Pareto optimal solution. Genetic algorithm and the probabilistic search characteristics and cluster analysis for the classification of the uncertainty of the results of unity. Therefore, the algorithm uses k-means clustering analysis to achieve group diversification, the specific steps are as follows:

- In Popsize individuals in groups, calculation of the various individual objective function value, according to each individual target function values randomly selected q individuals, produces q Center (q < Popsize), and the center point of the corresponding numerical value as the center of gravity;
- In the remaining Popsize-q individual take an individual, calculate it to q center distance, and distribute it to a recent center;
- Calculation of the center of gravity, as the new center of gravity;
- By analogy, until the Popsize-q individual is assigned, and q cluster groups.

In this algorithm, the feasible Pareto groups clustering analysis operation, generating q cluster groups, each in a cluster group of randomly selected individuals, they will be combined into a group, we call this group for clustering optimal Pareto groups, Pareto groups that feasible can be divided into Pareto groups clustering and the clustering of Pareto optimal population.
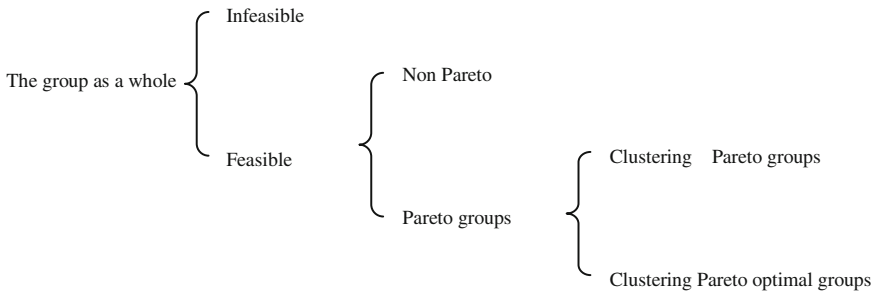
So that the entire group eventually classified into four groups, namely: infeasible groups, viable non-Pareto groups, clustering Pareto groups and clustering Pareto optimal groups. Relationships are as follows:

3.4 Group R-Fitness Assignment

In the above four groups, infeasible groups are in the entire group adaptation of the worst group, should reduce as they passed from one generation to the next group of probability; and the clustering of Pareto optimal population is from Pareto groups in carefully selected the most adaptive elite groups, should improve their inherited from one generation to the next group of probability. In order to achieve this goal, for these four groups were given R fitness, their relationship:

R(Infeasible groups) $\leq$ R(Feasible Non Pareto groups) $\leq$ R(Clustering Pareto groups) $\leq$ R(Clustering Pareto optimal population)

The R-fitness can be used as algorithm, it will be as the genetic algorithm to choose the basis of the operation.

3.5 Algorithm Specific Processes are as Follows

- Setting the evolution algebra t ← M 1, generating initial population consists of the initial individual.
- Based on the constraint conditions, the whole colony is divided into feasible group $p(t)_1$ and infeasible group $p(t)_1'$.
- For feasible group $p(t)_1'$, calculation of the various individual adaptive value, then carries on the Pareto operation, feasible group divided into viable non Pareto group $p(t)_2$ and available Pareto group $p(t)_2'$.
- For feasible Pareto group $p(t)_2'$, k- means clustering analysis algorithm, the feasible Pareto group is divided into Pareto group $p(t)_3$ and Pareto clustering optimal group $p(t)_4$, which is the realization of the diverse groups an important step.
- On $p(t)_1$, $p(t)_2$, $p(t)_3$, $p(t)_4$, the four sub groups, respectively, given proper fitness value.
- The group of proportional selection operation.
- The group of single cross operation.
- The groups were uniform mutation operation.
- Termination condition. If the termination condition is not satisfied, then update the evolution algebra counter t ← 2, if the termination condition is satisfied, then the output calculation results, the algorithm ends.

4 Test results

In order to test the performance of the algorithm, using Matlab [11] programming, to two groups with complex constrained multi-objective optimization problems are calculated. Specific parameters for: population size *PopSize* = 200, gene length

ChroLength = 10, crossover probability $p_c = 0.7$, mutation probability $p_m = 0.1$, R (infeasible goup) = 1, R(Practical non Pareto group) = 20, R (cluster Pareto group) = 20, R (cluster Pareto optimal group) = 100.

Problem 1

$f_1(x) = (x_1 - 2)^2 + (x_2 - 1)^2 + 2$

$f_2(x) = 9x_1 - (x_2 - 1)^2$

subject to

$$x_1^2 + x_2^2 - 225 \leq 0$$

$$x_1 - 3x_2 + 10 \leq 0$$

$$-20 \leq x_1, x_2 \leq 20$$

Problem 2

$f_1(x) = 4x_1^2 + 4x_2^2$

$f_2(x) = (x_1 - 5)^2 + (x_2 - 5)^2$

subject to

$$(x_1 - 5)^2 + x_2^2 - 25 \leq 0$$

$$-(x_1 - 8)^2 - (x_2 + 3)^2 + 7.7 \leq 0$$

$$-15 \leq x_1, x_2 \leq 30$$

problem 1 optimal solutions as shown in Fig. 2, problem 2 optimal solutions as shown in Fig. 3:

From Figs. 2 and 3 can see, using the algorithm presented by this paper, not only the calculated Pareto optimal quantity, wide distribution, uniform, and the optimal frontier curve is smooth, shapely, in addition, the proposed algorithm evolutionary rate is rapidder, Fig. 2 only 10 generations of evolution, Fig. 3 evolution that has reached the 40 generation a good optimization effect.

Fig. 2 Problem 1 optimal solutions
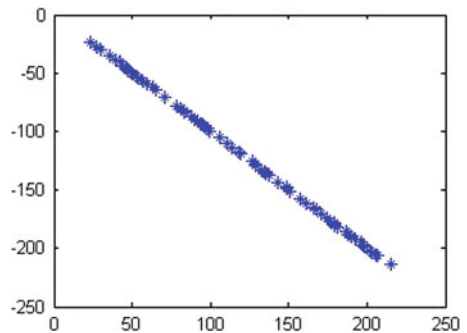
Fig. 3 Problem 2 optimal solutions
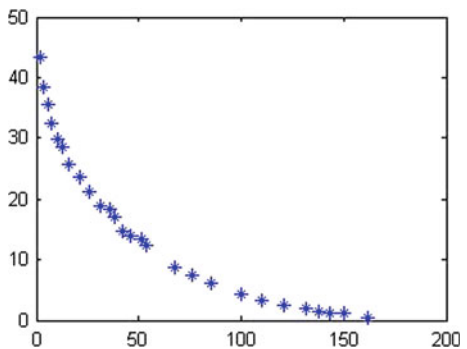
5 Postscript

This paper presents a method based on group classification of complex constrained multi-objective optimization genetic algorithm [12], the algorithm for the solution of some highly constrained optimization problem to provide the effective way, it uses clustering and classification concept, greatly improves the performance of the algorithm, improving the understanding of assembly quality. Computer simulation results show that, the algorithm can not only be widely distributed, even Pareto optimal solution, and the evolutionary speed, usually only 10 $\sim$ 40 generation can achieve very good results.

References

1. Schaffer JD (1985) Multiple objective optimization with vector evaluated genetic algorithms. In: Grefenstette J (ed) Proceedings of an international conference on genetic algorithms and their applications
2. Horn J, Nafpliotis N, Goldberg DE (1994) A niched Pareto genetic algorithm for multiobjective optimization. In: Proceedings of the first IEEE conference on evolutionary computation. Piscataway, pp 82–87
3. Fonseca CM, Fleming PJ (1993) Genetic algorithms for multiobjective optimization: formulation, discussion and generalization. In: Proceeding of the fifth international conference on genetic algorithms. Morgan Kauffman, San Francisco
4. Deb K, Agrawal S, Pratap A, Meyarivan T (200) A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II, KanGAL report No. 200001
5. Zitzler E, Deb K, Thiele L (2000) Comparison of multiobjective evolutionary algorithms: empirical results.Evol Comput 8(2):125–147
6. Jimenez F, Verdegay J (1999) Evolutionary techniques for constrained optimization problems. In: 7th European congress on intelligent techniques and soft computing (EUFIT'99), Springer, Aachen, Germany
7. Wang Y, Liu L (2005) Constrained multi-objective optimization evolutionary algorithm. J Tsinghua Univ (Nat Sci Ed) 45(1):103–106
8. Wang X, Cao L (2002) Genetic algorithm theory, application and software implementation. Xi'an Jiao Tong University press, Xi'an
9. Wen X, Zhou L, Wang D (2000) MATLAB Neural network application design. Science press, Beijing
10. Wang L (2001) Intelligent optimization algorithm and its application. Qinghua University press, Beijing
11. Xing W, Xie J (2000) Modern optimization method. Qinghua University press, Beijing
12. Anderson R, Lew D (2003) Evaluating predictive models of species distributions: criteria for selecting optimal models. Ecol Model 162(3):211–232

Hypothesis Testing for Squared Radial Ornstein–Uhleneck Model: Moderate Deviations Method

Cai-Ping Chen, Shou-Jiang Zhao and Qiao-Jing Liu

Abstract We study the moderate deviations for the log-likelihood ratio of the squared radial Ornstein-Uhleneck (O–U) model, with the help of them we give negative regions in testing squared radial O–U model, and get the decay rates of the error probabilities.

Keywords Hypothesis testing · Moderate deviations · Squared radial Ornstein-Uhleneck model

1 Introduction

We consider the hypothesis testing problem for the following squared radial Ornstein-Uhleneck model:

$$dX_t = (\delta + 2bt)dt + 2\sqrt{X_t}dW_t, \quad X_0 = 1 \tag{1}$$

where $\delta > 2$ is the unknown parameter to be tested on the basis of continuous observation of the process $\{X_t, t \geq 0\}$ on the time interval $[0, T]$, W is a standard Brownian motion, and $b \leq 0$ is known. We denote by P_b^δ the distribution of the solution (1).

We decide the two hypothesis:

$$H_0 : \delta = \delta_0 \quad \text{and} \quad H_1 : \delta = \delta_1$$

where $\delta_0, \delta_0 > 2$. The hypothesis testing is based on a partition of Ω_T of the outcome process on $[0, T]$ into two (decision) regions B_T and its compliment B_T^C,

C.-P. Chen · S.-J. Zhao (✉) · Q.-J. Liu
China Three Gorges University, College of Science, 443002 Yichang,
People's Republic China
e-mail: shjzhao@yahoo.com.cn

171

and we decide that H_0 is true or false according to the outcome $X \in B_T$ or $X \in B_T^C$. By the Neyman-Pearson lemma [1], the optional decision region B_T has the following form:

$$\left\{ \log \frac{dP_b^{\delta_1}}{dP_b^{\delta_0}} \big|_{F_T} \geq c_T \right\}$$

The research of hypothesis testing problem has started in the 1930s. Since the optional decision region B_T has the above form, we are interested in the calculation or approximation of the constant c_T, and the hypothesis testing problem can be studied by large deviations [2–8]. In those papers, some large deviation estimates of the error probabilities for some i.i.d. sequences, Markov chains, stationary Gaussian processes, Ornstein-Uhlenbeck processes, Jacobi model are obtained. In this paper, we study the moderate deviations for the hypothesis testing problem of squared radial Ornstein-Uhlenck model, by moderate deviation principle of the log-likelihood ratio, we give negative regions and get the decay rates of the error probabilities. The large and moderate deviations for parameter estimators of squared radial Ornstein-Uhlenbeck model were studied in [9] and [10].

Throughout this paper, we assume: $a(T)$ be a positive function satisfying

$$\frac{a(T)}{T} \to 0 \quad \text{and} \quad \frac{a(T)}{\sqrt{T}} \to \infty \text{ as } T \to \infty$$

Now we state our main results.

Theorem 1 *If $b = 0$, for any $c > 0$, set*

$$B_T = \left\{ \frac{1}{a(\log T)} \log \frac{dP_0^{\delta_1}}{dP_0^{\delta_0}} \bigg|_{F_T} \geq \frac{|\delta_1^2 - \delta_0^2|}{2\delta_0} \sqrt{\frac{c}{2\delta_0}} - \frac{(\delta_1 - \delta_0)^2 \log T}{8\delta_0 a(\log T)} \right\}$$

Then

$$\lim_{T \to \infty} \frac{\log T}{a^2(\log T)} \log P_0^{\delta_0}(B_T) = -c$$

and

$$\lim_{T \to \infty} \frac{\log T}{a^2(\log T)} \log P_0^{\delta_1}(B_T^C) = -\infty.$$

Theorem 2 *If $b < 0$, for any $c > 0$, set*

$$B_T = \left\{ \frac{1}{a(T)} \log \frac{dP_b^{\delta_1}}{dP_b^{\delta_0}} \bigg|_{F_T} \geq \frac{|\delta_1^2 - \delta_0^2|}{2\delta_0} \sqrt{\frac{-cb}{\delta_0}} - \frac{(\delta_1 - \delta_0)^2 \log \rho(T)}{8\delta_0 a(T)} \right\}$$

Then

$$\lim_{T\to\infty}\frac{T}{a^2(T)}\log P_b^{\delta_0}(B_T)=-c$$

and

$$\lim_{T\to\infty}\frac{T}{a^2(T)}\log P_b^{\delta_1}(B_T^C)=-\infty,$$

where

$$\rho(T)=\frac{e^{-2bT}-1}{-2b}.$$

2 Moderate Deviations in Testing Squared Radial Ornstein-Uhleneck Model

In this section, we will prove Theorem 1 and 2. Let us introduce the log-likelihood ratio process of squared radial Ornstein-Uhlenbeck model, and study the moderate deviations of the log-likelihood ratio process. By [10], the log-likelihood ratio process has the representation

$$\log\frac{dP_b^{\delta_1}}{dP_b^{\delta_0}}|_{F_t}=\frac{\delta_1-\delta_0}{4}\int_0^t\frac{dX_s}{X_s}-\frac{\delta_1^2-\delta_0^2}{8}\int_0^t\frac{ds}{X_s}-\frac{b}{2}(\delta_1-\delta_0)t. \qquad (2)$$

The following two Lemmas [10, 11] plays an important role in this paper.

Lemma 3 *Let $\{X_t, t\geq 0\}$ be the process defined by* (1). *Then the following properties hold.*

1. *For $k>0$, the law of $\{k^{-1}X_{kt}, t\geq 0\}$ under $P_{kx,0}^{\delta}$ is $P_{x,0}^{\delta}$.*
2. *For $b<0$, the law of*

$$\left\{e^{2bt}X\left(\frac{e^{-2bt}-1}{-2b}\right),\ t\geq 0\right\}$$

under P_0^{δ} is P_b^{δ}.

Lemma 4 *If $b=0$, then $\{X_t, t\geq 0\}$ has a transition density function defined by*

$$p_t^{\delta}=\begin{cases}\frac{1}{2t}(y/x)^{v/2}\exp\{-\frac{x+y}{2t}\}I_v(\sqrt{xy}/t) & x>0, y\in R^+\\[2mm]
(2t)^{-\delta/2}\Gamma(\delta/2)^{-1}y^v e^{-\frac{y}{2t}} & x=0, y\in R^+\end{cases}$$

where I_v is the Bessel function of index v, and $v = \frac{\delta-2}{2}$. If $b = 0$, we consider the moderate deviations for the log-likelihood ratio process of squared radial Ornstein-Uhleneck model.

Lemma 5 *For any closed subset $F \subset R$,*

$$\limsup_{T \to \infty} \frac{\log T}{a^2(\log T)} \log P_0^{\delta_0} \left(\frac{1}{a(\log T)} \left[\log \frac{dP_0^{\delta_1}}{dP_0^{\delta_0}} |_{F_T} + \frac{(\delta_1 - \delta_0)^2 \log T}{8\delta_0} \right] \in F \right)$$

$$\leq - \inf_{z \in F} \frac{8\delta_0^3 z^2}{\left(\delta_1^2 - \delta_0^2\right)^2}$$

and for any open subset $G \subset R$,

$$\liminf_{T \to \infty} \frac{\log T}{a^2(\log T)} \log P_0^{\delta_0} \left(\frac{1}{a(\log T)} \left[\log \frac{dP_0^{\delta_1}}{dP_0^{\delta_0}} |_{F_T} + \frac{(\delta_1 - \delta_0)^2 \log T}{8\delta_0} \right] \in G \right)$$

$$\geq - \inf_{z \in G} \frac{8\delta_0^3 z^2}{\left(\delta_1^2 - \delta_0^2\right)^2}$$

Proof Let

$$\Lambda_T(y) = \log E_0^{\delta_0} \exp \left\{ \frac{a(\log T)y}{\log T} \left[\log \frac{dP_0^{\delta_1}}{dP_0^{\delta_0}} |_{F_T} + \frac{(\delta_1 - \delta_0)^2 \log T}{8\delta_0} \right] \right\}.$$

By (2), for any $\varphi > 0$, we have

$$\Lambda_T(y) = \log E_0^{\delta_0} \exp \left\{ \frac{a(\log T)y}{\log T} \left[\log \frac{dP_0^{\delta_1}}{dP_0^{\delta_0}} |_{F_T} + \frac{(\delta_1 - \delta_0)^2 \log T}{8\delta_0} \right] \right\} \log \frac{dP_0^{\delta_0}}{dP_0^{\varphi}} |_{F_T}$$

$$= \log E_0^{\varphi} \exp \left\{ \left(\lambda + \frac{\delta_0 - \varphi}{4} \right) \log X_T - \left(u + \frac{\delta_0^2 - \varphi^2}{8} \right) \int_0^T \frac{ds}{X_s} - \lambda \log T + \frac{\log T}{\delta_0} u \right\}$$

where

$$\lambda = \frac{a(\log T)(\delta_1 - \delta_0)}{4 \log T} y$$

and

$$u = \frac{a(\log T)(\delta_1^2 - \delta_0^2)}{8 \log T} y.$$

For T large enough, we can choose $\varphi = \sqrt{\delta_0^2 + 8u}$, such that $8u + \delta_0^2 - \varphi^2 = 0$.

By Lemma 3, we get

$$E_{1,0}^{\varphi}\left(X_T^{\lambda+\frac{\delta_0-\varphi}{4}}\right) = T^{\lambda+\frac{\delta_0-\varphi}{4}}E_{1/T,0}^{\varphi}\left(X_1^{\lambda+\frac{\delta_0-\varphi}{4}}\right)$$

and

$$\Lambda_T(y) = \log E_{1/T,0}^{\varphi}\left(X_1^{\lambda+\frac{\delta_0-\varphi}{4}}\right) + \frac{\delta_0-\varphi}{4}\log T + \frac{\log T}{\delta_0}u.$$

Since for $p > 0$ (cf. [12])

$$\lim_{x\to 0^+}\frac{I_p(x)}{x^p} = 2^p p!,$$

therefore, by Lemma 4 and Taylor formula,

$$\begin{aligned}\Lambda_T(y) = &: \lim_{T\to\infty}\frac{\log T}{a^2(\log T)}\Lambda_T(y)\\ = &\lim_{T\to\infty}\frac{\log T}{a^2(\log T)}\left[\log E_{1/T,0}^{\varphi}\left(X_1^{\lambda+\frac{\delta_0-\varphi}{4}}\right) + \frac{a^2(\log T)y^2(\delta_1^2-\delta_0^2)^2}{32\delta_0^3\log T} + o\left(\frac{a^2(\log T)}{\log T}\right)\right]\\ = &\frac{(\delta_1^2-\delta_0^2)^2}{32\delta_0^3}y^2.\end{aligned}$$

Finally, the Gartner-Ellis theorem [13] implies the conclusion of lemma 5. Noting that

$$\log\frac{dP_0^{\delta_1}}{dP_0^{\delta_0}}\bigg|_{F_T} = -\log\frac{dP_0^{\delta_0}}{dP_0^{\delta_1}}\bigg|_{F_T},$$

we also have the following result.

Lemma 6 *For any closed subset $F \subset R$,*

$$\begin{aligned}&\limsup_{T\to\infty}\frac{\log T}{a^2(\log T)}\log P_0^{\delta_1}\left(\frac{1}{a(\log T)}\left[\log\frac{dP_0^{\delta_1}}{dP_0^{\delta_0}}|_{F_T} - \frac{(\delta_1-\delta_0)^2\log T}{8\delta_1}\right] \in F\right)\\ &\leq -\inf_{z\in F}\frac{8\delta_1^3 z^2}{(\delta_1^2-\delta_0^2)^2}\end{aligned}$$

and for any open subset $G \subset R$,

$$\begin{aligned}&\liminf_{T\to\infty}\frac{\log T}{a^2(\log T)}\log P_0^{\delta_1}\left(\frac{1}{a(\log T)}\left[\log\frac{dP_0^{\delta_1}}{dP_0^{\delta_0}}|_{F_T} - \frac{(\delta_1-\delta_0)^2\log T}{8\delta_1}\right] \in G\right)\\ &\geq -\inf_{z\in G}\frac{8\delta_1^3 z^2}{(\delta_1^2-\delta_0^2)^2}\end{aligned}$$

If $b < 0$, we get the following two results for the log-likelihood ratio process of squared radial Ornstein-Uhleneck model.

Lemma 7 *For any closed subset $F \subset R$,*

$$\limsup_{T \to \infty} \frac{T}{a^2(T)} \log P_b^{\delta_0} \left(\frac{1}{a(T)} \left[\log \frac{dP_b^{\delta_1}}{dP_b^{\delta_0}} \Big|_{F_T} + \frac{(\delta_1 - \delta_0)^2 \log \rho(T)}{8\delta_0} \right] \in F \right)$$

$$\leq - \inf_{z \in F} \frac{4\delta_0^3 z^2}{-b(\delta_1^2 - \delta_0^2)^2}$$

and for any open subset $G \subset R$,

$$\liminf_{T \to \infty} \frac{T}{a^2(T)} \log P_b^{\delta_0} \left(\frac{1}{a(T)} \left[\log \frac{dP_b^{\delta_1}}{dP_b^{\delta_0}} \Big|_{F_T} + \frac{(\delta_1 - \delta_0)^2 \log \rho(T)}{8\delta_0} \right] \in G \right)$$

$$\geq - \inf_{z \in G} \frac{4\delta_0^3 z^2}{-b(\delta_1^2 - \delta_0^2)^2}$$

where $\rho(T) = \frac{e^{-2bT}-1}{-2b}$.

Proof Let

$$\Lambda_T(y) = \log E_b^{\delta_0} \exp\left\{ \frac{a(T)y}{T} \left[\log \frac{dP_b^{\delta_1}}{dP_b^{\delta_0}} \Big|_{F_T} + \frac{(\delta_1 - \delta_0)^2 \log \rho(T)}{8\delta_0} \right] \right\}.$$

By (2), for any $\varphi > 0$, we have

$$\Lambda_T(y) = \log E_b^{\delta_0} \exp\left\{ \frac{a(T)y}{T} \left[\log \frac{dP_b^{\delta_1}}{dP_b^{\delta_0}} \Big|_{F_T} + \frac{(\delta_1 - \delta_0)^2 \log \rho(T)}{8\delta_0} \right] \right\} \log \frac{dP_b^{\delta_0}}{dP_b^{\varphi}} \Big|_{F_T}$$

$$= \log E_b^{\varphi_0} \exp\left\{ \left(\lambda + \frac{\delta_0 - \varphi}{4} \right) \log X_T - \left(u + \frac{\delta_0^2 - \varphi^2}{8} \right) \int_0^T \frac{ds}{X_s} \right.$$

$$\left. - \lambda \log \rho(T) - 2bT\lambda - \frac{b}{2}(\delta_0 - \varphi)T + \frac{\log \rho(T)}{\delta_0} u \right\}$$

where

$$\lambda = \frac{a(T)(\delta_1 - \delta_0)}{4T} y$$

and

$$u = \frac{a(T)(\delta_1^2 - \delta_0^2)}{8T} y.$$

For T large enough, we can choose $\varphi = \sqrt{\delta_0^2 + 8u}$, such that $8u + \delta_0^2 - \varphi^2 = 0$.

By Lemma 3, we get

$$E_{1,b}^{\varphi}\left(X_T^{\lambda+\frac{\delta_0-\varphi}{4}}\right) = \exp\left\{2bT\left(\lambda+\frac{\delta_0-\varphi}{4}\right)\right\}\rho(T)^{\lambda+\frac{\delta_0-\varphi}{4}}E_{1/T,0}^{\varphi}\left(X_1^{\lambda+\frac{\delta_0-\varphi}{4}}\right),$$

where $\rho(T) = \frac{e^{-2bT}-1}{-2b}$.

And then

$$\Lambda_T(y) = \log E_{1/T,0}^{\varphi}\left(X_1^{\lambda+\frac{\delta_0-\varphi}{4}}\right) + \frac{\delta_0-\varphi}{4}\log\rho(T) + \frac{\log\rho(T)}{\delta_0}u.$$

Applying the Taylor formula to the function $\sqrt{1+x}$, we have
$\sqrt{\delta_0^2+8u} = \delta_0\left[1 + \frac{a(T)y(\delta_1^2-\delta_0^2)}{2\delta_0^2 T} - \frac{a^2(T)y^2(\delta_1^2-\delta_0^2)^2}{8\delta_0^4 T^2} + o\left(\frac{a^2(T)}{T}\right)\right].$

Therefore, by Lemma 4,

$$\Lambda_T(y) =: \lim_{T\to\infty}\frac{T}{a^2(T)}\Lambda_T(y)$$

$$= \lim_{T\to\infty}\frac{T}{a^2(T)}\left[\log E_{1/T,0}^{\varphi}\left(X_1^{\lambda+\frac{\delta_0-\varphi}{4}}\right) + \frac{a^2(T)y^2(\delta_1^2-\delta_0^2)^2}{32\delta_0^3 T^2}\log\rho(T) + o\left(\frac{a^2(T)}{T}\right)\right],$$

$$= -\frac{b(\delta_1^2-\delta_0^2)^2}{16\delta_0^3}y^2$$

since $\log\rho(T)/T \to -2b$, as $T \to \infty$. Now, we can apply the Gartner-Ellis theorem to get the conclusion of lemma 7.

We also have the following result using the same method as in lemma 7.

Lemma 8 *For any closed subset $F \subset R$,*

$$\limsup_{T\to\infty}\frac{T}{a^2(T)}\log P_b^{\delta_1}\left(\frac{1}{a(T)}\left[\log\frac{dP_b^{\delta_1}}{dP_b^{\delta_0}}|_{F_T} - \frac{(\delta_1-\delta_0)^2\log\rho(T)}{8\delta_1}\right] \in F\right)$$

$$\leq -\inf_{z\in F}\frac{4\delta_1^3 z^2}{-b(\delta_1^2-\delta_0^2)^2}$$

and for any open subset $G \subset R$,

$$\liminf_{T\to\infty}\frac{T}{a^2(T)}\log P_b^{\delta 1}\left(\frac{1}{a(T)}\left[\log\frac{dP_b^{\delta_1}}{dP_b^{\delta_0}}|_{F_T} - \frac{(\delta_1-\delta_0)^2\log\rho(T)}{8\delta_1}\right] \in G\right)$$

$$\geq -\inf_{z\in G}\frac{4\delta_1^3 z^2}{-b(\delta_1^2-\delta_0^2)^2}$$

3 Proof of Main Theorem

Since the proofs of the two theorems are similar, we only prove Theorem 1. The first claim is a direct conclusion of Lemma 5. Since

$$-\frac{\log T}{a(\log T)}\frac{(\delta_1 - \delta_0)^2}{8\delta_0} - \frac{\log T}{a(\log T)}\frac{(\delta_1 - \delta_0)^2}{8\delta_1} \to -\infty, \ as \ T \to \infty$$

by Lemma 6, we see that the second one also holds.

References

1. Neyman J, Pearson ES (1933) On the problem of the most efficient tests of statistical hypotheses. Phil Trans Roy Soc London 231:289–337
2. Blahut RE (1984) Hypothesis testing and information theory. IEEE Trans Inform Theory 20:405–415
3. Chiyonobu T (2003) Hypothesis testing for signal detection problem and large deviations. Nagoya Math J 162:187–203
4. Gao FQ, Zhao SJ (2012) Moderate deviations and hypothesis testing for signal detection problem. Sci China Math 55:2273–2284
5. Han TS, Kobayashi K (1989) The strong converse theorem in hypothesis testing. IEEE Trans Inform Theory 35:178–180
6. Jiang H, Zhao SJ (2011) Large and moderate deviations in testing time inhomogeneous diffusions. J Stat Plan Infer 141:3160–3169
7. Nakagawa K, Kanaya F (1933) On the converse theorem in statistical hypothesis testing for Markov chains. IEEE Trans Inform Theory 39:629–633
8. Zhao SJ, Gao FQ (2010) Large deviations in testing Jacobi model. Stat Probab Lett 80:34–41
9. Gao FQ, Jiang H (2009) Moderate deviations for squared radial Ornstein-Uhlenbeck process. Stat Probab Lett 79:1378–1386
10. Zani M (2002) Large deviations for squared radial Ornstein-Uhlenbeck processes. Stoch Process Appl 102:25–42
11. Watson GN (1995) A treatise on the theory of bessel function. Cambridge University Press Cambridge
12. Pitman J, Yor M (1982) A decomposition of Bessel bridges. Z Wahrscheinlichkeitstheor Verwandte Geb 59:425–457
13. Dembo A, Zeitouni O (1988) Large deviation technique and applications. Springer, New York

A Nonconforming Characteristic Finite Element Method for Nonlinear Advection-Dominated Diffusion Equation with Memory Term

Jiaquan Zhou, Chao Xu and Jianlai Gao

Abstract A nonconforming characteristic finite element method is considered for nonlinear convection-dominated diffusion equation with memory term. By the use of some special properties of the finite element interpolation operator, and without Rietz-Volterra projection operator which is an indispensable tool in the convergence analysis of finite element methods for integro-differential evolution equations in the previous literature, the optimal error estimate on L^2-norm and the superconvergence result on H^1-norm are obtained.

Keywords Nonlinear convection-dominated diffusion equation with memory $\cdot$ Nonconforming characteristic finite element $\cdot$ The optimal error estimate $\cdot$ Superconvergence

1 Introduction

Consider the following nonlinear convection-dominated diffusion equation with memory term

J. Zhou (✉) · C. Xu · J. Gao
Department of Mathematics and Physics, Luoyang Institute of Science and Technology, Luoyang 471023, People Republic of China
e-mail: lyzhjq@126.com

C. Xu
e-mail: xc-lyct@126.com

J. Gao
e-mail: lygaojianlai@163.com

$$
\begin{cases}
c(x,u)\frac{\partial u}{\partial t} + d(x,u) \cdot \nabla u - \nabla \cdot [a(x,u)\nabla u] \\[2mm]
\quad + \int_0^t b(x,u(\tau))\nabla u(\tau)d\tau] = f(x,u), & (x,t) \in \Omega \times (0,T], \\[2mm]
u(x,t) = 0, & (x,t) \in \partial\Omega \times (0,T], \\[2mm]
u(x,0) = u_0(x), & x \in \Omega,
\end{cases}
\tag{1}
$$

where $\Omega \subset R^2$ is a bounded convex polygonal domain with Lipschitz continuous boundary $\partial\Omega$, $(0,T]$ is the time interval, $u_0(x)$ and $f(x,u)$ is are given function and smooth function respectively. Assume that

A_1 There exist constants $a_0, a_1, b_0, b_1, c_0, c_1$, such that $0 < \delta_0 \leq \delta(x,u) \leq \delta_1 < +\infty$, $\delta = a, b, c$.

A_2 $a(x,u)$, $b(x,u)$ and $f(x,u)$ are global Lipschitz continuous with respect to u. i.e., there exists positive constant L such that

$$
|\varphi(u_1) - \varphi(u_2)| \leq L|u_1 - u_2|, \ \varphi = a, b, f.
$$

A_3 $d(x,u) = (d_1(x,u), d_2(x,u))$, $c(x,u), d(x,u)$ are Lipschitz continuous with respect to u. and there is a positive constant K_1 such that

$$
|d(x,u)| + |\nabla d(x,u)| + |\frac{d(x,u)}{c(x,u)}| + |\nabla \cdot \frac{d(x,u)}{c(x,u)}| \leq K_1.
$$

A_4 The initial value $u_0 \in H^2(\Omega) \cap H_0^1(\Omega)$, the true solution u is existent and unique, and satisfies

$$
u \in L^\infty\left(0,T;W^{2,\infty}(\Omega) \cap H^3(\Omega) \cap H_0^1(\Omega)\right), \frac{\partial u}{\partial t} \in L^2\left(0,T;H^2(\Omega)\right),
$$

and exists a positive constant K_2 such that $||u||_{L^\infty(L^\infty)} \leq K_2$.

It is well known that strictly parabolic discretization schemes applied to the problem do not work well when it is advection dominated. It is especially difficult to approximate well the sharp fronts and to conserve the material or mass in the system.

When the general finite element methods or finite difference procedures are used to solve the convection-dominated diffusion problem, it exhibits excessive numerical diffusion and nonphysical oscillations [1, 2]. To overcome the above difficulties, [3] suggested characteristic finite element method for the linear convection-dominated problem. Compared with the standard methods, this method not only has a simple scheme, and reduces the time-truncation error, because of its solution changing more slowly in the characteristic direction than in the time direction, but also allows to use the large time steps without losing accuracy. On the other hand, for the domain decomposition methods, since the unknowns are associated with the element faces, and each degree of freedom belongs to at most two elements, the use of the Crouzeix-Raviart type nonconforming finite elements facilitates the exchange of information across each subdomain and provides

spectral radius estimates for the iterative domain decomposition operator. This results in cheap local communication, and the method can be parallelized in a highly efficient manner on MIMD-machines [4].

In this paper, we propose a nonconforming characteristic element scheme for nonlinear convection-dominated diffusion equation with memory term. By virtue of some special properties of the finite element interpolation operator, and without Rietz-Volterra projection operator which is an indispensable tool in the convergence analysis of finite element methods for integro-differential evolution equations in the previous literature, the optimal error estimate on L^2-norm and the superconvergence result on H^1-norm are obtained.

2 Construction of Nonconforming Finite Element

Let $\Omega \subset R^2$ be a polygon with boundaries $\partial\Omega$ parallelling the axes, T_h be a partition of Ω by rectangular meshes, satisfying the usual regularity condition or quasi-uniform assumption. $\forall K \in T_h$, let $K = (x_K - h_{x,K}, x_K + h_{x,K}) \times (y_K - h_{y,K}, y_K + h_{y,K})$, the four vertices of K are $a_1(x_K - h_x, y_K - h_y)$, $a_2(x_K + h_x, y_K - h_y)$, $a_3(x_K + h_x, , y_K + h_y)$, $a_4(x_K - h_x, y_K + h_y)$, and the four edges are $l_1 = \overline{a_1 a_2}$, $l_2 = \overline{a_2 a_3}$ $l_3 = \overline{a_3 a_4}$, $l_4 = \overline{a_4 a_1}$, and the length of edges parallel to x-axis and y-axis by $2h_{x,K}$, $2h_{y,K}$ respectively. Let $h_K = \mathrm{diam}(K)$, $h_K = \max\limits_{K \in T_h} h_K$.

The shape function spaces P and interpolation operators Π of the finite elements on K are defined by

$$P = \mathrm{span}\{1, x, y, \varphi(x), \varphi(y)\};$$

$$\frac{1}{|l_j|}\int_{l_j}(v - \Pi v)ds = 0, \quad (j = 1, 2, 3, 4), \quad \frac{1}{|K|}\int_K (v - \Pi v)dxdy = 0,$$

where $\varphi(t) = \frac{1}{2}(3t^2 - 1)$.

Then the associated finite element spaces V_h can be defined as

$$V_h = \left\{ v : v|_K \in P, \forall K \in T_h, \int_l [v]ds = 0, l \subset \partial K \right\}.$$

where $[v]$ denotes the jump of v across the boundary l if l is an internal edge, and $[v] = v$ if $l \subset \partial\Omega$.

Obviously, $V_h \not\subset H^1(\Omega)$, so V_h is a nonconforming finite element space.

Then for all $v \in H^1(\Omega)$ let Π_h be the associated interpolation operators on V_h.

$$\Pi_h : H^1(\Omega) \rightarrow V_h, \Pi_h|_K = \Pi_K, \Pi_K v = (\Pi v).$$

We define mesh dependent inner product over V_h as $(u_h, v_h)_h = \sum\limits_{K \in T_h} \int_K u_h \cdot v_h dx$, and it is easy to see that $||u_h||^2_{0,h} = \sum\limits_{K \in T_h} ||u_h||^2_{0,K}$ and $||u_h||^2_h = \sum\limits_{K \in T_h} |u_h|^2_{1,K}$ are the norms over V_h. Without causing confusion, $(u_h, v_h)_h$ is still written as (u_h, v_h).

For $u \in H^1(\Omega)$, $\forall v \in V_h$, by employing the similar arguments, we can get

$$(\nabla(u - \Pi_h u), \nabla v) = 0. \tag{2}$$

Here and later, the positive constant C will be used as a generic constant, which is independent of the mesh parameter h.

In this paper, let $W^{m,p}(\Omega)$ be the standard Sobolev space with the usual norm $|| \cdot ||_k$ and semi-norm $| \cdot |_k$ respectively. When $k = 0$, we let $L^2(\Omega)$ denote the corresponding space defined on Ω with norm $|| \cdot ||$. The following spaces that incorporate time dependence will be used. Let $[a, b] \subset (0, T]$, Y be a Sobolev space, and $f(x, t)$ be smooth function on $\Omega \times [a, b]$, also we define $L^p(a, b; Y)$ and $||f||_{L^p(a,b;Y)}$, as $L^p(a, b; Y) = \left\{ f : \int_a^b ||f(\cdot, t)||^p_Y dt < \infty \right\}$ and $||f||_{L^p(a,b;Y)} = \left(\int_a^b ||f(\cdot, t)||^p_Y dt \right)^{\frac{1}{p}}$, if $p = \infty$, the integral is replaced by the essential supremum.

3 Construction of Nonconforming Characteristics Finite Element Scheme

Let $\psi(u) = \sqrt{c^2(x, u) + |d(x, u)|^2}$ and denote the characteristic direction associated with the operator $c(x, u)\frac{\partial u}{\partial t} + d(x, u) \cdot \nabla u$ as $\tau = \tau(u)$, where

$$\frac{\partial}{\partial \tau} = \frac{c(x, u)}{\psi(x, u)} \frac{\partial}{\partial t} + \frac{d(x, u)}{\psi(x, u)} u \cdot \nabla$$

Then (1) can be put in the form

$$\begin{cases} \psi(x, u)\frac{\partial u}{\partial \tau} - \nabla \cdot [a(x, u)\nabla u] \\ \quad + \int_0^t b(x, u(\tau))\nabla u(\tau)d\tau] = f(x, u), & (x, t) \in \Omega \times (0, T], \\ u(x, t) = 0, & (x, t) \in \partial\Omega \times (0, T], \\ u(x, 0) = u_0(x), & x \in \Omega, \end{cases} \tag{3}$$

The weak form of (3) is: find $u \in H_0^1(\Omega) \cap H^2(\Omega)$ such that

$$
\begin{cases}
(\psi(x,u)\frac{\partial u}{\partial \tau}, v) - (a(x,u)\nabla u) \\
\quad + \int_0^t b(x,u(\tau))\nabla u(\tau)d\tau, \nabla v) = (f(x,u),v), \quad \forall v \in H_0^1(\Omega), \\
u(x,0) = u_0(x), \qquad\qquad\qquad\qquad\qquad x \in \Omega,
\end{cases}
\tag{4}
$$

Let a time step $\Delta t > 0$, $N = \frac{T}{\Delta t}$ be a positive integer, and approximate the solution at times $t^n = n\Delta t$, and the characteristic derivative will be approximated basically in the following manner.

Let $\overline{X} = X - \frac{d(x,u^{n-1})}{c(x,u^{n-1})}\Delta t$, we have the following approximation similar to [3]

$$
\left[\psi(x,u)\frac{\partial u}{\partial \tau}\right](x,t^n) \approx \psi(x,u^n)\frac{u(x,t^n) - u(\overline{x},t^{n-1})}{\left[(x-\overline{x})^2 + (\Delta t)^2\right]} = c(x,u^n)\frac{u(x,t^n) - u(\overline{x},t^{n-1})}{\Delta t}
$$

$$\tag{5}$$

And for giving smooth functions $f(\tau)$ and $g(\tau)$, we have

$$
\int_{t^i}^{t^{i+1}} f(\tau)g(\tau)d\tau = \Delta t f(t^i)g(t^i) + \varepsilon_i(f,g), \; \|\sum_{i=0}^n \varepsilon_i(f,g)\| = O(\Delta t).
$$

Then the characteristic nonconforming finite element scheme of (4) is: find $u_h : \{t^1, t^2, \cdots, t^N\} \in V_h$, such that

$$
\begin{cases}
(c(u_h^n)\frac{u_h^n - \overline{u}_h^{n-1}}{\Delta t}, v_h)_h + (a(u_h^n)\nabla u_h^n, \nabla v_h)_h \\
\quad + \sum_{i=0}^n (b(u_h^i)\nabla u_h^i, \nabla v_h)_h = (f(u_h^n), v_h)_h, \quad \forall v_h \in V_h \\
u_h^0 = \Pi_h u_0, \qquad\qquad\qquad\qquad\qquad \forall x \in \Omega,
\end{cases}
\tag{6}
$$

where, $u_h^n = u_h(x,t^n)$, $\overline{u}_h^{n-1} = u_h(\overline{x}, t^{n-1})$, $c(u_h^n) = c(x, u_h^n)$, $f(u_h^n) = f(x, u_h^n)$, $\Pi_h u_0$ is finite element interpolation of $u_0(x)$.

We see that the algebraic problems at each time step are linear and the matrix in (6) is independent of time for the constant time steps being employed. So the solution of (6) is unique.

4 The Error Estimates

At first, we present the following lemmas which are crucial to our error analysis.

Lemma 1 [5–7] *For any $u \in H^2(\Omega), \varphi \in V_h$, we have*

$$||u - \Pi_h u||_{0,h} + h||u - \Pi_h u||_h \le Ch^2 |u|_2, \qquad (\text{and}7)$$

$$|\sum_{K \in J_h} \int_{\partial K} \frac{\partial u}{\partial \boldsymbol{n}} \phi ds| \le Ch|u|_2 ||\varphi||_h. \qquad (8)$$

Furthermore, if $u \in H^3(\Omega)$, then

$$||u - \Pi_h u||_h \le Ch^3 |u|_3, \qquad (9)$$

and

$$|\sum_{K \in J_h} \int_{\partial K} \frac{\partial u}{\partial \boldsymbol{n}} \phi ds| \le Ch^2 |u|_3 ||\phi||_h. \qquad (10)$$

Here and later, $\boldsymbol{n}$ denotes the unit outer normal vector on K.

Lemma 2 *Assume that* $u \in L^\infty(H^1(\Omega) \cap W^{2,\infty}(\Omega)). \forall t \in (0, T]$, *then* $\nabla \Pi_h u(t)$ *and* $\Pi_h u(t)$ *are uniformly bounded, i.e.,*

$$||\nabla \Pi_h u(t)||_{0,h,\infty} + ||\Pi_h u(t)||_{0,h,\infty} \le C. \qquad (10)$$

where $||\phi(t)||_{0,h,\infty} = \max_K ||\phi(t)||_{0,K,\infty}, \forall \phi \in V_h, t \in (0, T]$.

Proof Let $P_0 u = \frac{1}{|K|} \int_K u dx$, using the triangle inequality and the inverse inequality, we have

$$||\Pi_h u||_{0,h,\infty} \le ||\Pi_h u - P_0 u||_{0,h,\infty} + ||P_0 u - u||_{0,h,\infty} + ||u||_{0,\infty}$$
$$\le Ch^{-1} \left(||\Pi_h u - P_0 u||_0 + ||P_0 u - u||_0 \right) + ||u||_{0,\infty}$$
$$\le C \left(||u||_1 + ||u||_{0,\infty} \right).$$
$$||\nabla \Pi_h u(t)||_{0,h,\infty} \le ||\nabla(\Pi_h u - u)||_{0,h,\infty} + ||\nabla u||_{0,h,\infty}$$
$$\le Ch||u||_{2,\infty} + ||\nabla u||_{0,h,\infty} \le C.$$

The uniform boundedness of $\nabla \Pi_h u(t)$ and $\Pi_h u(t)$ are proven.

Lemma 3 [3] *Let* $\eta \in L^2(\Omega)$, *and* $\overline{\eta} = \eta(x - g(x)\Delta t)$, *where function g and its gradient* ∇g *are bounded, then*

$$||\eta - \overline{\eta}||_{-1} \le C\Delta t ||\eta||.$$

The following theorems are the error estimate results of nonconforming characteristic finite element scheme (6).

Theorem 1 *Let u and* u_h *denote the solutions of* (4) *and* (6) *respectively. Then under assumptions* $\mathbf{A}_1 \sim \mathbf{A}_4$, *for sufficiently small* $\Delta t > 0$, *we have*

$$\max_{0 \le n \le N} ||(u - u_h)(t^n)|| \le m_0 \Delta t + m_1 h^2. \tag{12}$$

where,

$$m_0 = C \left\| \frac{\partial^2 u}{\partial \tau^2} \right\|_{L^2(0,T;L^2)} + 1, \; m_1 = C \left(|u_t|_{L^2(0,T;H^2)} + |u|_{L^\infty(0,T;H^2)} + |u|_{L^\infty(0,T;H^3)} \right).$$

Proof Let $u_h - u = (u_h - \Pi_h u) - (u - \Pi_h u) = \theta - \rho$. By (4) and (6), and choosing $v_h = \theta^n$. we obtain the error equation

$$
\left(c(u_h^n) \frac{\theta^n - \overline{\theta}^{n-1}}{\Delta t}, v_h \right) + \left(a(u_h^n) \nabla \theta^n, \nabla \theta^n \right)
$$

$$
= \left(\psi(u^n) \frac{\partial u^n}{\partial \tau} - c(u^n) \frac{u_h^n - \overline{u}_h^{n-1}}{\Delta t}, \theta^n \right) - \left(c(u_h^n) \frac{\rho^n - \overline{\rho}^{n-1}}{\Delta t}, \theta^n \right)
$$

$$
+ \left(\frac{c(u^n) - c(u_h^n)}{\Delta t} \Pi_h u^n, \theta^n \right) + \left(\frac{c(u_h^n) - c(u^n)}{\Delta t} \Pi_h \overline{u}^{n-1}, \theta^n \right)
$$

$$
+ \left(a(u^n) \nabla \rho^n, \nabla \theta^n \right) + \left(\Delta t \sum_{i=0}^{n-1} b(u_h^i) \nabla (\rho^i - \theta^i), \nabla \theta^n \right)
$$

$$
+ \left((a(u^n) - a(u_h^n)) \nabla \Pi_h u^n, \nabla \theta^n \right) + \left(\sum_{i=0}^{n-1} \varepsilon_i (b(u^i), \nabla u^i), \nabla \theta^n \right)
$$

$$
+ \left(\Delta t \sum_{i=0}^{n-1} (b(u^i) - b(u_h^i)) \nabla \Pi_h u^i, \nabla \theta^n \right) + \left(f(u_h^n) - f(u^n), \theta^n \right)
$$

$$
+ \sum_{K \in J_h} \int_{\partial K} a(u^n) \frac{\partial u^n}{\partial n} \theta^n ds + \Delta t \sum_{i=0}^{n-1} \sum_{K \in J_h} \int_{\partial K} b(u_h^i) \frac{\partial u^i}{\partial n} \theta^n ds
$$

$$
= \sum_{k=0}^{12} I_k. \tag{13}
$$

For the first term I_1 on the right-hand side of (13), by Cauchy-Schwartz inequality, ε-Young inequality, and using the same arguments as [8], we get

$$|I_1| \le C \Delta t \left\| \frac{\partial^2 u}{\partial \tau^2} \right\|_{L^2(0,T;L^2)} + \varepsilon ||\theta^n||^2. \tag{14}$$

For the terms $I_2 \sim I_{12}$ on the right-hand side of (13), imposing Cauchy inequality, assumption A_1, A_3, Lemma 1, Lemma 2, Lemma 3 and ε-Young inequality, we have

$$|I_2| \le C \frac{1}{\Delta t} ||\rho_t||_{L^2(0,T;L^2)} + C ||\rho^{n-1}||^2 + \varepsilon ||\theta^n||^2 + \varepsilon ||\theta^n||_h^2. \tag{15}$$

$$|I_3| + |I_4| \le C \frac{1}{\Delta t} \left(||\theta^n||^2 + ||\rho^n||^2 \right) + \varepsilon ||\theta^n||^2. \tag{16}$$

$$|I_5| + |I_6| \le C||\nabla\rho^n||^2 + C(\Delta t)^2 \left(\sum_{i=0}^{n-1} ||\nabla\rho^i||^2 + \sum_{i=0}^{n-1} ||\nabla\theta^i||^2 \right) + \varepsilon||\nabla\theta^n||^2. \quad (17)$$

$$\sum_{i=7}^{10} |I_i| \le C\left(||\rho^n||^2 + ||\theta^n||^2 + (\Delta t)^2 \right) + C(\Delta t)^2 \sum_{i=0}^{n-1} \left(||\rho^i||^2 + ||\theta^i||^2 \right) \quad (18)$$
$$+ \varepsilon||\theta^n||^2 + \varepsilon||\nabla\theta^n||^2$$

$$|I_{11}| + |I_{12}| \le Ch^4|u^n|_3^2 + Ch^4(\Delta t)^2 \sum_{i=0}^{n-1} |u^i|_3^2 + \varepsilon||\theta^n||_h^2. \quad (19)$$

Noting that $\int_\Omega \overline{\theta}^{n-1}\theta^n dx \le \frac{1}{2}\left[\int_\Omega \left(\overline{\theta}^{n-1} \right)^2 dx + \int_\Omega (\theta^n)^2 dx \right]$ and $\left(\overline{\theta}^{n-1}, \overline{\theta}^{n-1} \right) \le (1 + C\Delta t)\left(\theta^{n-1}, \theta^{n-1} \right)$ the left-hand side of (43) can be bounded by

$$\left(c(u_h^n)\frac{\theta^n - \overline{\theta}^{n-1}}{\Delta t}, v_h \right) + \left(a(u_h^n)\nabla\theta^n, \nabla\theta^n \right)$$
$$\ge \frac{c_0}{2\Delta t}\left[||\theta^n||^2 - (1 + C\Delta t)||\theta^{n-1}||^2 \right] + \frac{a_0}{2}||\nabla\theta^n||^2. \quad (20)$$

Therefore, combining (14–20) with (13), multiplied by $2\Delta t$, summing them in time, and applying Gronwall lemma, we obtain

$$||\theta^n|| \le m_0\Delta t + m_1 h^2. \quad (21)$$

Note that $u_h - u = \theta - \rho$, by Lemma 1 and the triangle inequality, we can get (12). This completes our first conclusion. By (9) and (10), and choosing $v_h = \theta^n - \theta^{n-1}$, similarly to the proof of Theorem 1, we can obtain the superclose results of H^1-norm as follows.

Theorem 2 *Let u and u_h denote the solutions of (4) and (6) respectively, and satisfy assumptions* $\mathbf{A}_1 \sim \mathbf{A}_4$, *for sufficiently small Δt, we have*

$$\max_{0 \le n \le N} ||(u - u_h)(t^n)||_h \le m_0\Delta t + m_1 h^2. \quad (22)$$

Furthermore, by using the same techniques as [9], the postprocessing interpolation operator Π_{2h} can be constructed, and satisfies the following properties (cf. [6])

$$\Pi_{2h}\Pi_h u = \Pi_{2h}u, ||\Pi_{2h}u - u||_h \le Ch^2|u|_3, ||I_{2h}v||_h \le C||v||_h, \forall v \in H^3(\Omega), v \in V_h. \quad (23)$$

Assume that u and u_h denote the solutions of (4) and (6) respectively, for all $u \in H_0^1(\Omega) \cap H^3(\Omega)$, by use of the similar arguments as [6], we can derive the following global superconvergence result.

Theorem 3 *Suppose that u and u_h denote the solutions of (4) and (6) respectively, and $u, u_t \in H^3(\Omega)$, we have the following superconvergence result*

$$||u^n - \Pi_{2h}u_h^n||_h \leq C\big(\Delta t + h^2\big). \tag{24}$$

References

1. Chen ZX, Chou SH, Kwak DY (2006) Characteristic-mixed covolume methods for advection-dominated diffusion problems. Numer Linear Algebra Appl 13:667–697
2. Arbogast T, Wheeler MF (1995) A characteristics-mixed finite element method for advection-dominated transport problems. SIAM J Numer Anal 32:404–424
3. Douglas J Jr, Russell TF (1982) Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures. SIAM J Numer Anal 19:871–885
4. Shi DY, Ren JC (2010) A least squares Galerkin-Petrov nonconforming mixed finite element method for the stationary conduction-convection problem. Nonlinear Anal: TMA 72:1653–1667
5. Shi DY, Mao SP, Chen SC (2005) An anisotropic nonconforming finite element with some superconvergence results. J Comput Math 23:261–274
6. Shi DY, Wang HH, Du YP (2009) An anisotropic nonconforming finite element method for approximating a class of nonlinear Sobolev equtions. J Comput Math 27:299–314
7. Ciarlet PG (1978) The finite element method for elliptic problem. North-Holland Pual, Amsterdam
8. Yuan YR (1993) Characteristic finite element method and analysis for the semiconductor device numerical simulation (in Chinese). Acta Math Sci 13A:241–251
9. Lin Q, Tobiska L, Zhou AH (2005) Superconvergence and extrapolation of nonconforming low order elements applied to the Poisson equation. IMA J Numer Anal 25:160–181

Exact Solutions for the Compound KdV-Burgers Equation Using the Improved (G′/G)-Expansion Method

Rui Cao

Abstract Exact solutions of nonlinear partial differential equation are major subject in the study of nonlinear science. In this paper, with the aid of computer algebraic system mathematica, the compound KdV-Burgers equation are investigated by using the improved (G′/G)-expansion method. As a result, many exact solutions are obtained which including hyperbolic function solutions, trigonometric function solutions and rational function solutions.

Keywords (G′/G)-expansion method · Compound KdV-Burgers equation · Exact solution

1 Introduction

The investigations of exact traveling wave solutions to nonlinear nonlinear partial differential equations (PDEs) play an important role in the study of nonlinear physical phenomena described by nonlinear PDEs. With the development of soliton theory, there have been many efficient methods available for seeking exact solutions of nonlinear PDEs, such as inverse scattering method, Backlund transformation, Hirota method, Lie group method, Darboux transformation, homogeneous balance method, various tanh function method, Riccati equation expansion method, Jacobi elliptic function method, F-expansion method and so on[1–10]. Abundant of the exact solutions, including the solitary wave solutions, the shock wave solutions, the periodic wave solutions and the soliton solutions, have been obtained by using these methods.

R. Cao (✉)
Department of Mathematics, Heze University, Heze 274015, China
e-mail: ruicao999@126.com

In Ref. [11], Wang has presented a straightforward method called as the (G′/G)-expansion method. By using the (G′/G)-expansion method, more exact traveling wave solutions of nonlinear PDEs are successfully obtained [12, 13]. The key idea of the method is that the traveling wave solutions of PDEs can be expressed by polynomial in $\varphi(\xi)$, where $\varphi(\xi) = \frac{G'(\xi)}{G(\xi)}$ and $G(\xi)$ satisfies the second order linear ordinary differential equation. Recently, much research work has been concentrated on the various extensions and applications of the (G′/G)-expansion method.

In this paper, our purpose is to construct exact solutions of compound KdV-Burgers equation by using the improved (G′/G)-expansion method. The rest of the paper is organized as follows. In Sect. 2, the improved (G′/G)-expansion method is introduced. In Sect. 3, we apply the proposed method to the compound Kdv-Burgers equation and many solutions are successfully obtained. Finally, some conclusions and discussions will be presented in Sect. 4.

2 Description of the Improved (G′/G)-Expansion Method

In the following we would like to outline the main steps of the improved (G′/G)-expansion method. For a given nonlinear PDE system with two variables x, t

$$P(x, t, u_t, u_x, u_{xx}, u_{tt}, u_{tx}, \cdots) = 0 \tag{2.1}$$

where $u = u(x, t)$ is an unknown function, the function P is a polynomial function of the field u and its partial derivatives in which the higher order derivatives and the nonlinear terms are involved.

Step 1. Using the following wave transformation

$$u(x, t) = u(\xi), \ \xi = kx + \omega t. \tag{2.2}$$

where k, ω are to be determined. Substituting (2.2) into (2.1) yields the ordinary differential equation (ODE) for $u(\xi)$

$$P(t, x, \omega u', ku', k^2 u'', \omega^2 u'', k\omega u'', \cdots) = 0. \tag{2.3}$$

Step 2. Supposing that the solution of (2.3) can be expressed as follows

$$u(\xi) = \sum_{i=-N}^{N} a_i \left(\frac{G'}{G}\right)^i, \tag{2.4}$$

where $G = G(\xi)$ satisfies the following second order linear ordinary differential equation

$$G''(\xi) + \mu G(\xi) = 0, \tag{2.5}$$

where a_i, μ are constants to be determined. The degree "N" can be determined by considering the homogeneous balance between the highest order derivatives and the nonlinear terms.

Step 3. Substituting (2.4) along with (2.5) into (2.3),collecting all terms with the same power of $\frac{G'(\xi)}{G(\xi)}$ together and then setting each coefficient of the resulted polynomial to zero yields a system of nonlinear algebraic equations for λ, μ, k, ω and a_i.

Step 4. Solving the algebraic equations obtained in the step 3 with the aid of Mathematica. Substituting the values of λ, μ, k, ω and a_i and the general solutions of (2.5) into (2.4) and (2.2), we can obtain more exact traveling wave solutions of the nonlinear PDEs (2.1).

3 Exact Solutions of the Compound KdV-Burgers Equation

We are concerned with the compound Kdv-Burgers equation in the following forms

$$u_t + puu_x + qu^2u_x + ru_{xx} - su_{xxx} = 0, \tag{3.1}$$

where p, q, r are constants. This system of equation can be seen as a compound equation of the KdV, mKdV and Burgers equations which involves nonlinear effect, dispersion effect and dissipation effect.

Using the traveling wave variables

$$u(x, t) = u(\xi), \ \xi = x - vt. \tag{3.2}$$

where v is a constant to be determined later. Eq. (3.2) can be converted into an ODE for $u = u(\xi)$

$$-vu' + puu' + qu^2u' + ru'' - su''' = 0, \tag{3.3}$$

Considering the homogeneous balance between the highest order derivatives and the nonlinear terms in Eq. (3.3), So we assume that Eq. (3.3) has solution in the following form

$$u(\xi) = a_{-1}\phi(\xi)^{-1} + a_0 + a_1\phi(\xi), \tag{3.4}$$

where a_{-1}, a_0, a_1 are constants to be determined and satisfy $a_{-1}^2 + a_1^2 \neq 0$, $\phi(\xi) = \frac{G'(\xi)}{G(\xi)}$ and $G(\xi)$ satisfies Eq. (2.5).

Substituting (3.4) into Eq. (3.3) and collecting all terms with the same power in ϕ, we can get a polynomial in ϕ. Setting each coefficient of $\phi^i (i = -4, -3, -2, -1, 0, 1, 2, 3, 4)$ to zero yields a system of algebraic equations as follows

$$q\mu a_{-1}^3 - 6s\mu^3 a_{-1} = 0,$$
$$p\mu a_{-1}^2 + 2q\mu a_{-1}^2 a_0 + 2r\mu^2 a_{-1} = 0,$$
$$-v\mu a_{-1} + p\mu a_0 a_{-1} + \mu q a_{-1}^2 a_1 + \mu q a_{-1} a_0^2 + q a_{-1}^3 - 8s\mu^2 a_{-1} = 0,$$
$$p a_{-1}^2 + 2\mu q a_0 a_1 a_{-1} + 2r\mu a_{-1} = 0,$$
$$-v(a_{-1} - \mu a_1) + p a_0 (a_{-1} - \mu a_1) + q(a_{-1} a_1 + a_0^2)(a_{-1} - \mu a_1) - 2\mu s(a_{-1} - \mu a_1) = 0,$$
$$-p\mu a_1^2 - 2\mu q a_0 a_1^2 + 2r\mu a_1 = 0,$$
$$v a_1 - p a_1 a_0 + q a_{-1} a_1^2 - \mu q a_1^3 - 2q a_{-1} a_1^2 - q a_0^2 a_1 + 8\mu s a_1 = 0,$$
$$p a_1^2 + 2q a_0 a_1^2 - 2r a_1 = 0,$$
$$q a_1^3 - 6s a_1 = 0.$$

Solving the above algebraic equations yields the following results

Case 1 $qs > 0$,

$$a_{-1} = 0, \ a_0 = -\frac{p}{2q} \pm \frac{r}{\sqrt{6qs}}, \ a_1 = \pm\sqrt{\frac{6s}{q}}, \ v = -\frac{p^2}{4q} + \frac{r^2}{6s} - 2s\mu.$$

Case 2 $qs > 0$,

$$a_{-1} = \pm\sqrt{\frac{6s\mu^2}{q}}, \ a_0 = -\frac{p}{2q} \pm \frac{r}{\sqrt{6qs}}, \ a_1 = 0, \ v = -\frac{p^2}{4q} + \frac{r^2}{6s} - 2s\mu.$$

Case 3 $qs > 0$,

$$a_{-1} = \pm\sqrt{\frac{6s\mu^2}{q}}, \ a_0 = -\frac{p}{2q} \pm \frac{r}{\sqrt{6qs}}, \ a_1 = \pm\sqrt{\frac{6s}{q}}, \ v = -\frac{p^2}{4q} + \frac{r^2}{6s} - 2s\mu.$$

According to Eq. (3.4) and the solution of Eq. (2.5) and cases 1–3, we deduce the following exact solutions.

Theorem 1 *Suppose that $qs > 0$ and $\mu < 0$. Then Eq. (3.1) admit exact solutions*

$$u_{11}(\xi) = -\frac{p}{2q} \pm \frac{r}{\sqrt{6qs}} \pm \sqrt{\frac{-6\mu s}{q}} \frac{A_1 \cosh\sqrt{-\mu}\xi + A_2 \sinh\sqrt{-\mu}\xi}{A_1 \sinh\sqrt{-\mu}\xi + A_2 \cosh\sqrt{-\mu}\xi},$$

$$u_{12}(\xi) = -\frac{p}{2q} \pm \frac{r}{\sqrt{6qs}} \pm \mu\sqrt{\frac{-6\mu^2 s}{q}} \frac{A_1 \sinh\sqrt{-\mu}\xi + A_2 \cosh\sqrt{-\mu}\xi}{A_1 \cosh\sqrt{-\mu}\xi + A_2 \sinh\sqrt{-\mu}\xi},$$

$$u_{13}(\xi) = \pm\sqrt{\frac{-6\mu^3 s}{q}} \left(\frac{A_1 \cosh\sqrt{-\mu}\xi + A_2 \sinh\sqrt{-\mu}\xi}{A_1 \sinh\sqrt{-\mu}\xi + A_2 \cosh\sqrt{-\mu}\xi}\right)^{-1} - \frac{p}{2q} \pm \frac{r}{\sqrt{6qs}}$$

$$\pm\sqrt{\frac{-6\mu s}{q}} \frac{A_1 \cosh\sqrt{-\mu}\xi + A_2 \sinh\sqrt{-\mu}\xi}{A_1 \sinh\sqrt{-\mu}\xi + A_2 \cosh\sqrt{-\mu}\xi},$$

Suppose the $qs > 0$ and $\mu > 0$. Then Eq. (3.1) admit exact solutions

$$u_{21}(\xi) = -\frac{p}{2q} \pm \frac{r}{\sqrt{6qs}} \pm \sqrt{\frac{6\mu s}{q}} \frac{A_1 \cos\sqrt{\mu}\xi - A_2 \sin\sqrt{\mu}\xi}{A_1 \sin\sqrt{\mu}\xi + A_2 \cos\sqrt{\mu}\xi},$$

$$u_{22}(\xi) = -\frac{p}{2q} \pm \frac{r}{\sqrt{6qs}} \pm \mu\sqrt{\frac{-6\mu^2 s}{q}} \left(\frac{A_1 \cos\sqrt{\mu}\xi - A_2 \sin\sqrt{\mu}\xi}{A_1 \sin\sqrt{\mu}\xi + A_2 \cos\sqrt{\mu}\xi}\right)^{-1},$$

$$u_{23}(\xi) = \pm\sqrt{\frac{6\mu^3 s}{q}} \left(\frac{A_1 \cos\sqrt{\mu}\xi - A_2 \sin\sqrt{\mu}\xi}{A_1 \sin\sqrt{\mu}\xi + A_2 \cos\sqrt{\mu}\xi}\right)^{-1} - \frac{p}{2q} \pm \frac{r}{\sqrt{6qs}}$$

$$\pm\sqrt{\frac{6\mu s}{q}} \frac{A_1 \cos\sqrt{\mu}\xi - A_2 \sin\sqrt{\mu}\xi}{A_1 \sin\sqrt{\mu}\xi + A_2 \cos\sqrt{\mu}\xi},$$

Suppose that $qs > 0$ and $\mu = 0$. Then Eq. (3.1) admit exact solutions

$$u_{31}(\xi) = -\frac{p}{2q} \pm \frac{r}{\sqrt{6qs}} \pm \sqrt{\frac{-6s}{q}} \frac{A_1}{A_1\xi + A_2},$$

$$u_{32}(\xi) = -\frac{p}{2q} \pm \frac{r}{\sqrt{6qs}} \pm \sqrt{\frac{-6\mu^2 s}{q}} \frac{A_1}{A_1\xi + A_2},$$

$$u_{33}(\xi) = \pm\sqrt{\frac{-6\mu^2 s}{q}} \left(\frac{A_1}{A_1\xi + A_2}\right)^{-1} - \frac{p}{2q} \pm \frac{r}{\sqrt{6qs}} \pm \sqrt{\frac{-6s}{q}} \frac{A_1}{A_1\xi + A_2},$$

4 Discussion and Conclusion

In summary, we investigate the KdV-Burgers equation by using the improved (G'/G)-expansion method and obtain the exact solutions of the KdV-Burgers equation, which have important physical meanings. Therefore the results maybe

useful in research of these equations in the field of physics. This method can also be applied to other nonlinear mathematical physics models. Further work on these aspect is worthy of studying.

References

1. Ablowitz MJ, Clarkson PA (1991) Solitons, nonlinear evolution equations and inverse scattering transform. Cambridge University Press, Cambridge
2. Hirota R (1980) Direct method of finding exact solutions of nonlinear evolution equations. In: Bullough R, Caudrey P (eds) Backlund transformation. Springer, Berlin
3. Hirota R, Satsuma J (1981) Soliton solutions of a coupled KdV equation. Phys Lett A 85:407–408
4. Olver PJ (1993) Applications of Lie groups to differential equations. Springer, New York
5. Matveev VB, Salle MA (1991) Darboux transformations and solitons. Springer, Berlin
6. Wang ML, Zhou YB, Li ZB (1996) Application of a homogeneous balance method to exact solutions of nonlinear equations in mathematical physics. Phys Lett A 216:67–75
7. Zarea SA (2009) The tanh method: A tool for solving some mathematical models, vol 41. Solitons and Fractals, Chaos, pp 979–988
8. Abdou MA (2007) The extended tanh method and its applications for solving nonlinear physical models. Appl Math Comput 190:988–996
9. Dai CQ, Zhang JF (2006) Jacobian elliptic function method for nonlinear differential-difference equations. Chaos, Solitons Fractals 27:1042–1047
10. Zhang S (2006) The periodic wave solutions for the (2+1)-dimensional Konopelchenko Dubrovsky Equations. Chaos, Solitons Fractals 30:1213–1220
11. Wang ML, Li XZ (2008) The (G'/G)-expansion method and traveling wave solutions of nonlinear evolution equations in mathematical physics. Phys Lett A 372:417–423
12. Bian CQ, Pang J, Jin LH, Ying XM (2010) Solving two fifth order strong nonlinear evolution equations by using the (G'/G)-expansion method. Commun Nonlinear Sci Numer Simul 15:2337–2343
13. Bekir A (2008) Application of the (G'/G)-expansion method for nonlinear evolution equations. Phys Lett A 372(19):3400–3406

The Design of Asynchronous Motor Performance Testing System Based on the MSP430f47187

Qing Wang, Weilong Li, Lihua Zhang and Sheng Lu

Abstract This thesis uses MSP430F47187 as the control core, through collecting the data of asynchronous motor no-load test and short-circuit test, and calculating the asynchronous motor performance parameters based on the circle diagram. Finally, we can obtain the visual performance test report in the operation terminal. Therefore, we can acknowledge the performance of the motor parameters intuitively and control the operation of the asynchronous motor better.

Keywords Asynchronous motor · MSP430F47187 · Performance test

1 Instruction

The traditional induction motor detection method is through testing on the scene, wiring a electromotor many times, recording the test data of asynchronous motor and calculating the test results manually, finally, we can have a comparative analysis of the data [1]. The disadvantage of this method is the big error and low precision caused in the process of calculating, the low test efficiency caused by repeated connection in the test. These disadvantages of traditional induction motor testing method is difficult to meet the requirements of the development of modern industry. Therefore, developing a kind of simple, quick asynchronous motor performance testing system has a important significance.

Q. Wang (✉) · S. Lu
Anhui Xinhua University, Huangshan, China
e-mail: xiaoqing5921@126.com

W. Li
The Science Research of Coal Institute of Anhui province, Hefei, China

L. Zhang
HeFei University of Technology, Hefei, China

The paper is based on the test characteristics of the asynchronous motor performance, combined with the performance testing requirements on the asynchronous motor of the current industry development. The paper designs the motor performance testing system based on the MSP430. The system with the MSP430F47187 as control core is based on the theory of the circle diagram, calculating the motor performance parameters by computer automatically by test of the asynchronous motor such as the no-load, short-circuit test and so on [2]. The design of the system consists of the machine design and the software design. The software design based on the VB6.0 software, designs various testing and reporting interface; And the machine design using the MSP430 as the hardware foundation, with the IAR as the software compiler environment, programming the data acquisition of the no-load test and short-circuit test and realizing the communication function of the machine [2].

2 The Design of the Hardware

2.1 The Block Diagram of the Hardware

The hardware system consists of the sampling system of the current and voltage, the communication system between the next-bit machine and the host computer, the printing system of the host computer, automatic pressure regulating system of the motor, the system circuit of the MSP430 microcontroller and so on. The voltage and current sensor collects the three-phase voltage and current by idling and stalling the motor, followed by the single-chip microcomputer transmitting the data to the host computer through the serial communication [3]. The host computer uses the previously compiled program to process the received data and obtain the performance parameters by the original computation from the motor. The structure diagram of the system is shown in Fig. 1.

Fig. 1 The structure diagram of the system
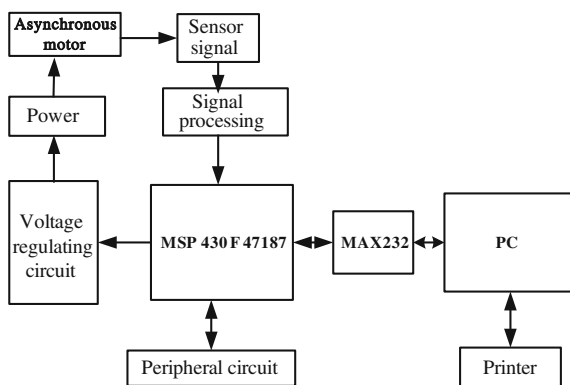
2.2 The Sampling Circuit

The sampling circuit of the testing system gets the desired voltage range through the amplification of the primary circuit, the two order filtering of the two stage circuit, and the voltage boosting of the three stage circuit. The sampling circuit diagram is shown in Fig. 2.

2.3 The Circuit of Voltage Regulating

The test system demands the motor realizing the no-load and locked-rotor test in different voltage levels. The system uses the MCU I/O port to control the relay switch which regulating the output voltage of the autotransformer (Autotransformer which has the public part of the transformer winding). As shown in Fig. 3, the single-chip machine drives the circuit to control the relay switch then control the switch of the contactor to control the autotransformer sequentially.

3 The Design of the Software

The software system consists of the lower position machine software and PC software, the lower position machine software system debugged the environment by the MCU software, compiled the sampling and control program of various parts by using C language and writes in the internal of the MCU. PC software system realizing the interface through the VB visualization. System flow diagram as shown in Fig. 4.
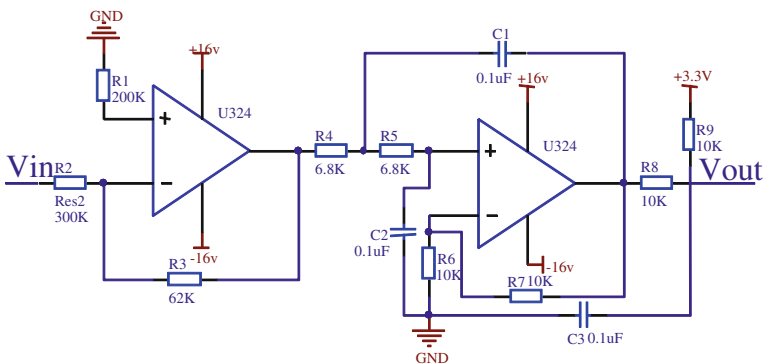
Fig. 2 The circuit of the sampling
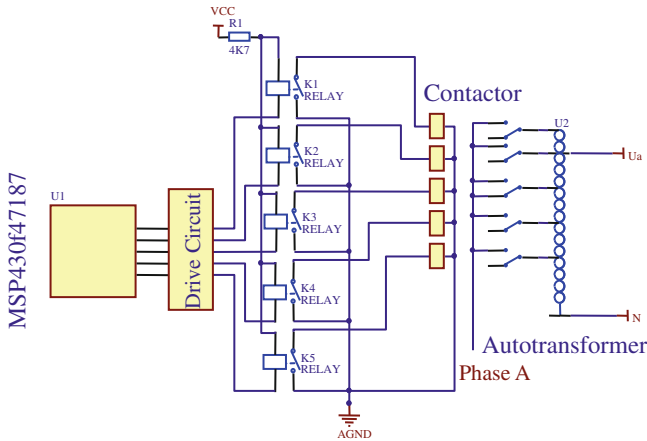
Fig. 3 The circuit of the voltage regulating

3.1 The Program Design of the Lower Position Machine

The program design of the lower position machine includes sampling, pressure regulating, communication etc. The sampling is the core program of the lower machine, need sampling the current or voltage value of 40 points in a waveform cycle, [4] calculating the root-mean-square value of the sampling data, divided by the number of sampling points, then we can get the effective value of the current or the voltage [5].

3.2 The Program Design of PC

PC program design mainly includes the following aspects:

1. The production of the basic parameter of motor interface, cold insulation resistance, direct current resistance interface, no-load test interface, locked rotor test interface, asynchronous motor report generation interface.
2. The production of the menu, menu is the drop-down menu hanging on the top of the interface, mainly used for the convenience of the jump, print and exit among different interfaces.
3. In the no-load and locked-rotor test, the design is used for transmitting the sent down start sampling to the lower position machine, and the communication of the lower position machine received by the PC software system.
4. The program processes the data transmitted by the lower position machine.
5. The program calculates the basic parameters of the motor with circle diagram method.

Fig. 4 The flow chart of the system

```
                              ┌──────────┐
                              │  Start   │
                              └──────────┘
                                   │
                                   ▼                          No
              ◇ Waiting for a host computer instruction ◇ ──────►
                                   │
                                  Yes
                                   │
                    ┌──────────────────────────────┐
                    │ The n =N (N is the number of  │
                    │ no-load voltage adjustment,   │
                    │ N=5, block N=2 )              │
                    └──────────────────────────────┘
                                   │
                    ┌──────────────────────────┐
                    │ Single chip microcomputer │      ┌──────────────────────────────┐
                    │ sampling and sending the  │      │ To adjust the output of the    │
                    │ data to PC                │      │ single chip computer I/O port  │
                    └──────────────────────────┘      │ control motor voltage          │
                                   │                   └──────────────────────────────┘
                                   ▼                             ▲
                    ◇       n=n-1,n=0?        ◇ ────────No────────┘
                                   │
                                  Yes
                                   │
                    ┌──────────────────────────┐
                    │ PC to accept data,        │
                    │ process data, fill in a   │
                    │ form                      │
                    └──────────────────────────┘
                                   │
                              ┌──────────┐
                              │   End    │
                              └──────────┘
```
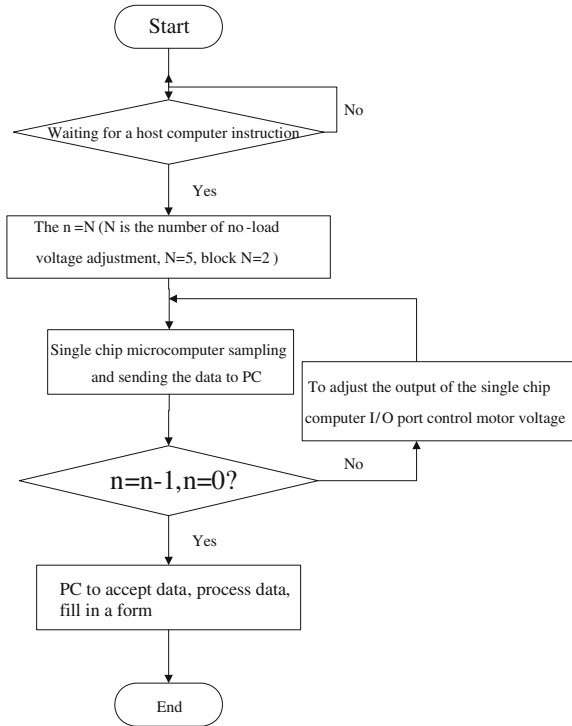
4 Debugging and Experiment

We debug the current test, voltage test successively, then adjust the power. Power calibration is the key step of the debugging, if we added a pure resistance load, so in this case, the product of the three-phase voltage and the three-phase current should be equal to the power values we calculated. Next we obtain the performance parameters of the motor by the no-load test and locked rotor test, finally we can obtain the test report. As shown in Fig. 5.

According to the test report, we can analyze the performance of the three-phase asynchronous motor, so we can judge whether the three-phase asynchronous motor is in accordance with the request for the continue use, if the motor can not continue to use, according to the actual situation, we may repair the parts of the motor that not meet the requirements.

5 Results

The test system with the MSP430F47187 as the control core, collected the data automatically, reduced the labor intensity of the testing work, improved the precision of the motor test results. The test system adopting the structure of the menu

Test Report of Three-Phase Asynchronous Motor

1. Original Data

Test Number	1	Unit	XFXY	Tester	Wang Qing	
Test Time	Rated Power	Rated Voltage	Rated Current	Rated Speed	Method	Insulation Class
2011-3-17	4000	380	8. 47	1440	Star	F or H

2. DC Resistance

Rab	Rbc	Rca	Ra	Rb	Rc	Rq
3. 1000	3. 1000	3. 1000	1. 5500	1. 5500	1. 5500	2. 3058

3. No-Load Experiment

Uab	Ubc	Uca	Ia	Ib	Ic	Po
259. 4462	260. 2193	261. 5423	5. 5342	5. 4613	5. 5512	460. 4521
242. 1235	243. 5312	241. 2315	4. 5284	4. 4512	4. 2145	288. 5632
220. 2151	220. 5641	220. 7546	3. 5642	3. 5478	3. 5421	212. 7845
No-Load Power	No-Load Current	Mechanical Loss	Iron Loss	No-Load Power Factor		
212. 5641	3. 5456	82. 0750	16. 3412	0. 1894		

4. Locked Rotor Experiment

Uab	Ubc	Uca	Ia	Ib	Ic	Po
50. 6645	49. 2548	50. 5463	9. 0123	9. 0011	9. 0423	1140. 8470
40. 8421	40. 5685	40. 2547	7. 2654	7. 2654	7. 2594	752. 6945
Rated Locked-Rotor Power	Rated Locked-Rotor Current	Locked-Rotor Power Factor	Locked-Rotor Torque			
7736. 5410	38. 7665	0. 6024	51308. 3000			

5. Experimental Result

Input Power	Input Current	Power Factor	Efficiency	Slip	Mechanical Loss	Copper Loss
42. 6731	8. 5234	0. 8450	0. 8200	0. 0414	82. 0750	74. 2634

Calculate

Print

Quit

Fig. 5 The test report of three phase asynchronous motor performance

added the interface operationally, is not only convenient and beautiful, but also practical. After the operation, the system is stable, reliable and can realize the performance test of the asynchronous motor quickly.

References

1. Wu Q (2010) The design of test line on small power motors factory. Electr Tools
2. Rao NSM (2011) Experience in short circuit testing of LV short circuit testing transformer. North American Power Symposium (NAPS) 1–4
3. Yang H, Zhou M (2012) The circuit test system based on MSP430. Electron Qual 7
4. Hu H (2004) Design and realization of DSP control system. Mechanical Industry Press, Beijing
5. Xie B, Ren Y Motor DSP control technology and its application. Aerospace, Beijing

Two-Step Iterative Algorithms for Two Multivalued Quasi-Nonexpansive Mappings

Yinying Zhou

Abstract In this paper, a new two-step iterative algorithms is introduced to approximate common Fixed point of two multivalued quasi-nonexpansive mappings in a real uniformly convex Banach Space. Furthermore, we also prove some strong and weak convergence theorems in uniformly convex Banach Space.

Keywords Nonexpansive mapping · Fixed point · Two-step iterative scheme

1 Introduction

Let X be a real Banach space. A subset D of X is called proximal, if for every $x \in X$, there exists an element $y \in D$, such that $d(x, y) = \inf\{\|x - z\|, z \in D\} = d(x, D)$.

It has been proved that weakly compact convex subsets of a Banach space and closed convex subsets of a uniformly convex Banach space are proximal. We shall denote the family of nonempty closed bounded subsets by $CB(X)$.

For $A, B \in CB(X)$, Let $H(A, B) = \left\{ \sup\limits_{x \in A} d(x, B), \sup\limits_{y \in B} d(y, A) \right\}$ (H be the Hausdorff metric induced by the metric of X and given by). A point $x \in D$ is called a fixed point of a multivalued mapping T, if $x \in T(x)$. We denote the set of all fixed points of T by $F(T)$. A multivalued mapping $T : D \to CB(X)$ is said to be

1. nonexpansive, if $H(Tx, Ty) \le \|x - y\|$ for every $x, y \in D$;
2. quasi nonexpansive, if $F(T) \ne \Phi$ and $H(Tx, Tp) \le \|x - p\|$ for every $p \in F(T)$.

Y. Zhou (✉)
Department of Mathematics and Information Sciences, Langfang Teacher's College,
Langfang 065000, China
e-mail: zhouyinying_hbu@163.com

In the sequel, we will write $F = F(S) \cap F(T)$ for the set of all common fixed points of the mappings S and T. Two multivalued nonexpansive mappings $T, S : D \to CB(X)$ are said to satisfy Condition (I), if there is a nondecreasing function $f : [0, +\infty) \to [0, +\infty)$ with $f(0) = 0, f(r) > 0$ for $r \in (0, +\infty)$ such that $d(x, Tx) \geq f(d(x, F)), d(x, Sx) \geq f(d(x, F)) (\forall x \in D)$

The theory of multivalued nonexpansive mapping is harder than the corresponding theory of single valued nonexpansive mappings. Different iterative processes have been used to approximate fixed points of multivalued nonexpansive mappings [1–5].

Sastry and Babu [1] considered the following algorithm. Let $T : X \to P(X)$ be a mapping with $P \in T(P)$, The sequence of Mann iterates is defined as

$$x = x_1 \in X, \ x_{n+1} = (1 - a_n)x_n + a_n y_n \tag{1}$$

where $y_n \in T(x_n)$, $\|y_n - p\| = d(T(x_n), p)$ and $\{a_n\}$ is a sequence in $(0, 1)$ satisfying $(0, 1)$. Danyanak improved the iteration scheme and gave the sequence of Ishikawa iterates as following: $x_1 = x \in X$

$$\begin{cases} y_n = (1 - b_n)x_n + b_n z_n, \ z_n \in T(x_n), \ \|z_n - p\| = d(T(x_n), p) \\ x_{n+1} = (1 - a_n)x_n + a_n u_n, \ u_n \in T(y_n), \ \|u_n - p\| = d(T(y_n), p) \end{cases} \tag{2}$$

$\{a_n\}, \{b_n\}$ are real sequences of numbers with $0 \leq a_n, b_n < 1$ satisfying $\lim\limits_{n \to \infty} b_n = 0$ and $\sum a_n b_n = \infty$.

On the other hand, Hu et al. [2] improved Ishikawa iterative as following: $x_1 = x \in X$

$$\begin{cases} y_n = (1 - b_n)x_n + b_n y_n, \ y_n \in T(x_n) \\ x_{n+1} = (1 - a_n)x_n + a_n u_n, \ u_n \in S(y_n), \ n \in N \end{cases} \tag{3}$$

where $\{a_n\}, \{b_n\}$ are in $[0, 1]$ satisfying certain conditions.

In this paper, a new Ishikawa iterative process is given to approximate common fixed points of two multivalued quasi-nonexpansive mapping.

Let D be a nonempty convex of a Banach space X and let $\{\alpha_n\}, \{\beta_n\},$ $\{\gamma_n\}; \{\alpha'_n\}, \{\beta'_n\}, \{\gamma'_n\} \subset (0, 1)$ with $\alpha_n + \beta_n + \gamma_n = 1, \ \alpha'_n + \beta'_n + \gamma'_n = 1,$ the sequence of Ishikawa iterates is defined by $x_0 \in D$

$$\begin{cases} y_n = \alpha'_n x_n + \beta'_n y''_n + \gamma'_n z''_n \\ x_{n+1} = \alpha_n x_n + \beta_n y'_n + \gamma_n z'_n \end{cases} \tag{4}$$

where $y''_n \in T(x_n), \ z''_n \in Sx_n, \ y'_n \in Ty_n, \ z'_n \in Sy_n$.

2 Preliminaries

We shall make use of the following result of Moha [3].

Lemma 2.1 *Let $R > 1$ be a fixed number, then X is a uniformly convex if and only if there exists a continuous strictly increasing and convex function $\varphi :$ $[0, +\infty) \rightarrow [0, +\infty)$ with $\varphi(0) = 0$ such that $\|\lambda x + (1 - \lambda)y\|^2 \leq \lambda\|x\|^2 + (1 - \lambda)\|y\|^2 - \lambda(1 - \lambda)\varphi(\|x - y\|)$ for all $x, y \in B_R(0) = \{x \in X : \|x\| \leq R\}$ and $\lambda \in [0, 1]$.*

3 Main Results

Lemma 3.1 *Let X be a uniformly convex Banach space, D be a nonempty bounded convex subset of X. $S, T : D \rightarrow CB(D)$ be a quasi-nonexpansive multi-valued map with $F = F(T) \cap F(S) \neq \Phi$ and for which $T(p) \cap S(\text{p}) = \{\text{p}\}$ for every $p \in F$, let $\{x_n\}$ be the Ishikawa iterates defined by (4), then $\lim_{n \to \infty}\|x_n - p\|$ exists for every $p \in F$.*

Proof By Equation (4), for every $p \in F$,

$$
\begin{aligned}
\|x_{n+1} - p\| &\leq \alpha_n\|x_n - p\| + \beta_n\|y'_n - p\| + \gamma_n\|z'_n - p\| \\
&= \alpha_n\|x_n - p\| + \beta_n d(y'_n, T(\text{p})) + \gamma_n d(z'_n, S(\text{p})) \\
&\leq \alpha_n\|x_n - p\| + \beta_n H(Ty_n, T(\text{p})) + \gamma_n H(Sy_n, S(\text{p})) \\
&\leq \alpha_n\|x_n - p\| + \beta_n\|y_n - p\| + \gamma_n\|y_n - p\| = \alpha_n\|x_n - p\| + (1 - \alpha_n)\|y_n - p\|
\end{aligned}
$$

$$
\begin{aligned}
\|y_n - p\| &\leq \alpha'_n\|x_n - p\| + \beta'_n\|y''_n - p\| + \gamma'_n\|z''_n - p\| \\
&\leq \alpha'_n\|x_n - p\| + \beta'_n d(y''_n, T(\text{p})) + \gamma'_n d(z''_n, S(\text{p})) \\
&\leq \alpha'_n\|x_n - p\| + \beta'_n H(Tx_n, T(\text{p})) + \gamma'_n H(Sx_n, S(\text{p})) \\
&\leq \alpha'_n\|x_n - p\| + \beta'_n\|x_n - p\| + \gamma'_n\|x_n - p\| \\
&= \|x_n - p\|
\end{aligned}
$$

Thus $\|x_{n+1} - p\| \leq \|x_n - p\|$.

Therefore, the sequence $\{\|x_n - p\|\}$ is decreasing and bounded below, thus $\lim_{n \to \infty}\|x_n - p\|$ exists for every $p \in F$ and $\{x_n\}$ is bounded.

Theorem 3.1 *Let X be a uniformly convex Banach space, D be a nonempty closed convex subset of X. $S, T : D \rightarrow CB(D)$ be a quasi-nonexpansive multi-valued map with $F \neq \Phi$ and $T(p) \cap S(\text{p}) = \{\text{p}\}$ for every $p \in F$, let $\{x_n\}$ be the Ishikawa iterates defined by (4), Assume that T satisfies condition (I) and $\{\alpha_n\}, \{\beta_n\}, \{\gamma_n\}; \{\alpha'_n\}, \{\beta'_n\}, \{\gamma'_n\} \subset [a, b] \subset (0, 1)$, then $\{x_n\}$ converges strongly to a common fixed point of S and T.*

Proof By Lemma 3.1, for each $p \in F$, $\{x_n\}$, $\{y_n\}$ is bounded. Therefore, there exists $R > 0$, such that $x_n - p$, $y_n - p \in B_R(0)$ for every $n \geq 0$. By Lemma 2.1, we have

$$\|x_{n+1} - p\|^2 \leq \alpha_n \|x_n - p\|^2 + \beta_n \|y'_n - p\|^2 + \gamma_n \|z'_n - p\|^2$$
$$- \alpha_n \beta_n \varphi(\|x_n - y'_n\|) - \alpha_n \gamma_n \varphi(\|x_n - z'_n\|) - \beta_n \gamma_n \varphi(\|y'_n - z'_n\|)$$
$$\leq \alpha_n \|x_n - p\|^2 + \beta_n [d(y'_n, Tp)]^2 + \gamma_n [d(z'_n, Sp)]^2$$
$$\leq \alpha_n \|x_n - p\|^2 + \beta_n [H(Ty_n, Tp)]^2 + \gamma_n [H(Sy_n, Sp)]^2$$
$$\leq \alpha_n \|x_n - p\|^2 + \beta_n \|y_n - p\|^2 + \gamma_n \|y_n - p\|^2 = \alpha_n \|x_n - p\|^2 + (1 - \alpha_n) \|y_n - p\|^2$$

$$\|y_n - p\|^2 \leq \alpha'_n \|x_n - p\|^2 + \beta'_n \|y''_n - p\|^2 + \gamma'_n \|z''_n - p\|^2$$
$$- \alpha'_n \beta'_n \varphi(\|x_n - y''_n\|) - \alpha'_n \gamma'_n \varphi(\|x_n - z''_n\|) - \gamma'_n \beta'_n \varphi(\|z'' - y''_n\|)$$
$$\leq \alpha'_n \|x_n - p\|^2 + \beta'_n [d(y''_n, Tp)]^2 + \gamma'_n [d(z''_n, Sp)]^2$$
$$- \alpha'_n \beta'_n \varphi(\|x_n - y''_n\|) - \alpha'_n \gamma'_n \varphi(\|x_n - z''_n\|) - \gamma'_n \beta'_n \varphi(\|z'' - y''_n\|)$$
$$\leq \alpha'_n \|x_n - p\|^2 + \beta'_n [H(Tx_n, Tp)]^2 + \gamma'_n [H(Sx_n, Sp)]^2$$
$$- \alpha'_n \beta'_n \varphi(\|x_n - y''_n\|) - \alpha'_n \gamma'_n \varphi(\|x_n - z''_n\|) - \gamma'_n \beta'_n \varphi(\|z'' - y''_n\|)$$
$$\leq \alpha'_n \|x_n - p\|^2 + \beta'_n \|x_n - p\|^2 + \gamma'_n \|x_n - p\|^2$$
$$- \alpha'_n \beta'_n \varphi(\|x_n - y''_n\|) - \alpha'_n \gamma'_n \varphi(\|x_n - z''_n\|) - \gamma'_n \beta'_n \varphi(\|z'' - y''_n\|)$$
$$= \|x_n - p\|^2 - \alpha'_n \beta'_n \varphi(\|x_n - y''_n\|) - \alpha'_n \gamma'_n \varphi(\|x_n - z''_n\|) - \gamma'_n \beta'_n \varphi(\|z'' - y''_n\|)$$

$$\|x_{n+1} - p\|^2 \leq \|x_n - p\|^2 - (\beta_n + \gamma_n) \alpha'_n \beta'_n \varphi(\|x_n - y''_n\|)$$
$$- (\beta_n + \gamma_n) \alpha'_n \gamma'_n \varphi(\|x_n - z''_n\|) - (\beta_n + \gamma_n) \gamma'_n \beta'_n \varphi(\|z'' - y''_n\|)$$

$$M\varphi(\|x_n - y''_n\|) \leq \|x_n - p\|^2 - \|x_{n+1} - p\|^2 \to 0, \text{ Let } M = (\beta_n + \gamma_n) \alpha'_n \beta'$$

Since $\sum M\varphi(\|x_n - y''_n\|) < +\infty$, we have $\lim_{n \to \infty} \varphi(\|x_n - y''_n\|) = 0$. Since φ is continuous at 0 and strictly increasing, we have $\lim_{n \to \infty} \|x_n - y''_n\| = 0$, thus $d(x_n, Tx_n) \to 0 (n \to \infty)$. Since T satisfies condition (I), $0 \leq f(d(x_n, F)) \leq d(x_n, Tx_n) \to 0$. Since f is strictly increasing, we have $\lim_{n \to \infty} d(x_n, F) = 0$, thus there exists a subsequence $\{x_{n_k}\} \subset \{x_n\}$ and a $\{p_k\} \subset F$, such that $\|x_{n_k} - p_k\| < \frac{1}{2^k}$ for all $k \geq 1$. By Lemma 3.1, we obtain $\|x_{n_{k+1}} - p_k\| \leq \|x_{n_k} - p_k\| < \frac{1}{2^k} \to 0 (k \to \infty)$, $\|p_{k+1} - p_k\| \leq \|x_{n_{k+1}} - p_{k+1}\| + \|x_{n_{k+1}} - p_k\| < \frac{1}{2^{k+1}} + \frac{1}{2^k} \to 0 (k \to \infty)$, Therefore $\{p_k\}$ is a cauchy sequence in D, let $\{p_k\}$ convergence to $q(k \to \infty)$. Since $0 \leq d(p_k, Tq) \leq H(Tp_k, Tq) \leq \|p_k - q\| \to 0$ and $p_k \to Tq$, $p_k \to q$, we have $d(q, Tq) = 0$ which implies that $q \in Tq$. Similarly $q \in Sq$. Since $\lim_{n \to \infty} \|x_n - p\|$ exists, $\{x_n\}$ converges strongly to $q \in F$.

References

1. Sastry KPR, Babu GVR (2005) Convergence of Ishikawa iterates for multivalued mapping with a fixed point. Czechoslovak Math J 55:817–826
2. Hu T, Huang JC, Rhoades BE (1997) A general principle for Ishikawa iterations for multivalued mappings. Indian J Pure Appl Math 28:1091–1098
3. Esla M, Abkar A (2011) One-step iterative process for a finite family of multivalued mappings. Math Comp Mode 54:105–111
4. Abbas M, Khan SH (2011) Common fixed points of two multivalued nonexpansive mappings by one-step iterative scheme. Appl Math Lett 24:97–102
5. Khan SH, Abbas M, Rhoades BE (2010) A new one-step iterative process for approximating common fixed points of two multivalued nonexpansive mappings. Rend Circ Mat 59:149–157

Analyzing Real-Time Predictability of Business Processes Based on Petri Nets

Wangyang Yu and Xianwen Fang

Abstract The real-time property is an important indicator of trustworthy business processes. Guaranteeing the real-time predictability is pivotal to ensure the efficient execution of business processes. For the real-time requirements of business processes, a method of modeling and analyzing business processes called Timed Business Process Net (TBPN) is proposed. It is a combination of business process modeled by Petri nets and the time factor on places. A real-time predictability analysis method is proposed. Using the proposed methods, we can provide an effective way to guarantee the real-time predictability of business processes.

Keywords Real-time · Predictability · Trustworthy · Business process

1 Introduction

Business process management (BPM) [1–3] is developing rapidly in today's business environment. How to efficiently construct a trustworthy business process to guarantee its ability to adapt quickly to new business requirements and competition is critical. Real-time property is an important indicator for trustworthy software system [4], and also for trustworthy business processes. Because most of the actual business processes have time limits. The time violation will increase business costs, so time performance analysis of business processes and

W. Yu (✉)
Department of Computer Science and Technology, Tongji University, Shanghai 201804, China
e-mail: ywy191@gmail.com

X. Fang
Faculty of Science, Anhui University of Science and Technology, Huainan 232001, China
e-mail: fangxianwen@hotmail.com

guaranteeing its real-time predictability is very important to enhance system management functions and system trustworthy.

Researches on business process management mainly focus on the reconstruction and optimization of business processes [5–10]. On the time management, existing works mainly focusing on workflow systems [11, 12] and WF_net [13, 14]. Petri nets are suitable to depict business processes as they are able to describe the systems in order, concurrent, conflict and synchronization [15]. For the business processes that require higher performance of real-time predictability, Timed Business Process Net (TBPN) is proposed in this paper. In a TBPN, places are used to represent the tasks. A task is associated with a time value. Transitions control the business logic and connect different tasks. A Token means a case in the business process. Optional tasks and alternative structures [11] are defined in TBPN.

Real-time predictability in a business process is that the system can execute according to the pre-planned time performance. To this end, we need to study the time plan of the business process, and try to avoid time violation for improving the efficiency of process management. Based on the operating rules and structural features of TBPN, a time calculation method is proposed, and time calculation and analysis will be done. Using the proposed method, we can calculate the time information effectively, so as it can provide the basic technical support for guarantee the real-time predictability of business processes.

2 Basic Concepts

Each task of the actual business process is represented by a place. The operation path framework shared by all cases is depicted by a TBPN model.

Definition 2.1 A Timed Business Process Net (TBPN) is a 6-tuple $\sum = (P, T; F, M, DI, ST)$, where

(1) $(P, T; F, M)$ an original Petri net;
(2) DI is a time function that is defined on S, i.e., $DI: S \rightarrow R_0$, R_0 is non-negative real number set;
(3) $P = S \cup S'$, S is the set of general task places, and S' is the set of optional tasks which are shown as double circle in order to represent the difference with S. $S \cap S' = \emptyset . ST : S' \rightarrow \{0\}$
(4) $F = F' \cup F''$, $\forall x, y \in P \cup T$, if $(x, y) \in F'$, draw a solid directed arc from x to y; Else if $(x, y) \in F''$, draw a dashed directed arc from x to y. F' and F'' are both flow relations, just expressed in different ways; and
(5) If $\exists x \in P$, for $\forall y \in x, (y, x) \in F'$, and for $\forall y' \in x, (x, y') \in F''$, then x is the beginning place of an alternative structure; If $\exists x' \in P$, for $\forall y \in x', (y, x') \in F''$, and for $\forall y' \in x', (x\prime, y') \in F'$, then x' is the end place of an alternative

structure. The subnet between the beginning place and the end place of an alternative structure is a S-graph, and the flow relations are dashed directed arcs.

For $t \in T, p \in S$, if $DI(p) = a, a \in R_0$, then, from the moment at which a token enters into p, p will not have the enabling function until the tokern stay within p at least a time units. If $t = \{p_1, p_2, \ldots, p_n\}, n \geq 1$, and $M[t>$, t will not fire until at least $max\{DI(p_1), DI(p_2), \ldots, DI(p_n)\}$ time units had passed. The firing is instantaneous.

A TBPN ia s Petri net extended with time factor, and constructed with some basic structures. For example, Fig. 1 is the schematic diagram of a optional task. In a TBPN, a place represents a task that is assigned a time value which means the execution delay of the task. Transitions connect tasks with each other, and they are used to reflect the logic relationships among tasks. In Fig. 1, $p_1(1)$, $p_2(2)$, $p_3(1)$ represent three tasks respectively. The figures in brackets indicate their execution delays. p_2 has a token, and this means the task is being executed. Its execution process consumes 2 time units. Figure 2 is the the basic composition structures of TBPN. The optional task can not be executed in case of an emergency [11]. The alternative tructure between p_1 and p_2 are two alternative paths which are choosen according to the the degree of tightness on time and policy information of the specific workflow instance. If permitted, selecting any path is correct.

Definition 2.2 A TBPN $\sum = (P, T; F, M, DI, ST)$ is structural correct if $(P, T; F, M)$ satisfies the following conditions:

(1) $\forall p \in P$, there exists a reachable marking $M \in R(M_0), M(p) \geq 1$; and
(2) $\forall t \in T$, there exists a reachable marking $M \in R(M_0), M[t >$.

Definition 2.2 guarantees that there are no dead tasks and dead transitions in a TBPN. We use reachable marking graph [15] to verify the structural correct of a business process.

3 Analysis Methods of Real-Time Predictability

Real-time predictability in a business process is that the system can execute according to the pre-planned time performance. First, we need to study the time plan of the business process. The execution delay of each task is allocated by the
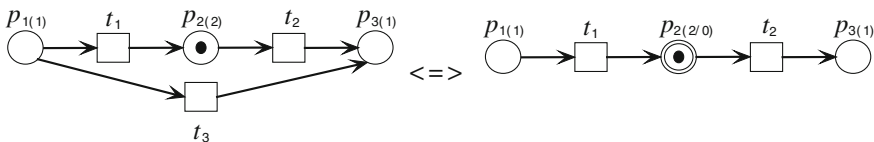
Fig. 1 Schematic diagram of an optional task
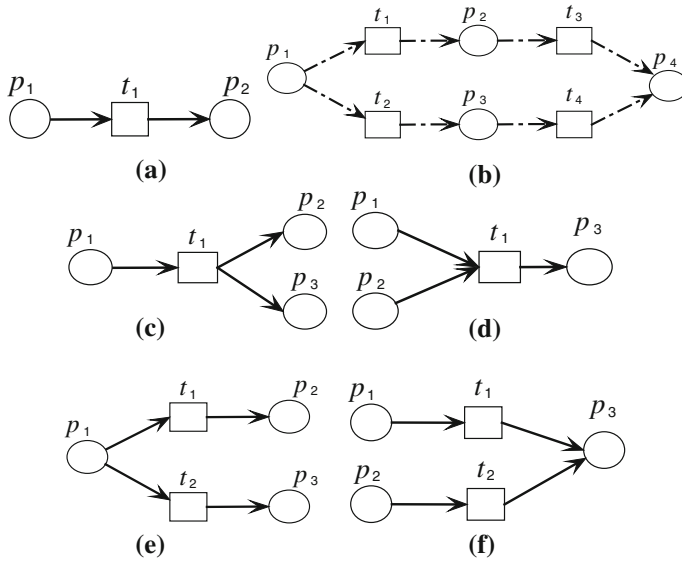
Fig. 2 Schematic diagram of basic composition structures. **a** Sequential structure. **b** Alternative structure. **c** Parallel branch structure. **d** Parallel converge structure. **e** Conditional branch structure. **f** Conditional converge structure

model designer, and the time value is either an average calculated by the previously executed multiple times or estimate value made by experts based on their experience and expectations. Here, based on the idea of [11], we define four basic paths among different tasks.

(1) *oac* path: Make optional tasks as the virtual task whose execution delay is 0; Choose the alternative path whose execution time is the least; At a conditional branch, according to the judgment of the actual external situation, choose the branch whose execution time is the least. The formed business path is called *oac* path.

(2) *õãc* path: Execute the optional tasks; Choose the alternative path whose execution time is the largest; At a conditional branch, according to the judgment of the actual external situation, choose the branch whose execution time is the least. The formed business path is called *õãc* path.

(3) *oac̃* path: Make optional tasks as the virtual task whose execution delay is 0; Choose the alternative path whose execution time is the least; At a conditional branch, according to the judgment of the actual external situation, choose the branch whose execution time is the largest. The formed business path is called *oac̃* path.

(4) *õãc̃* path: Execute the optional tasks; Choose the alternative path whose execution time is the largest; At a conditional branch, according to the judgment of the actual external situation, choose the branch whose execution time is the largest. The formed business path is called *õãc̃* path.

Then, $\forall p \in P$, we define the following time concepts according to the four basic paths:

(1) Γ_p^{oac}: The earliest time at which p can terminate according to oac path.
(2) $\Gamma_p^{\tilde{o}ac}$: The earliest time at which p can terminate according to $\tilde{o}ac$ path.
(3) $\Gamma_p^{oa\tilde{c}}$: The earliest time at which p can terminate according to $oa\tilde{c}$ path.
(4) $\Gamma_p^{\tilde{o}a\tilde{c}}$: The earliest time at which p can terminate according to $\tilde{o}a\tilde{c}$ path.

For a TBPN, the first work is to assign time delays to every $p \in P$. Then, a forward traversal is needed to compute Γ_p^{oac}, $\Gamma_p^{\tilde{o}ac}$, $\Gamma_p^{oa\tilde{c}}$, and $\Gamma_p^{\tilde{o}a\tilde{c}}$ for every $p \in P$. At the beginning, Γ values of the tasks without predecessor are set to 0. The formulas are as follow:

(1) If $p, j \in P$, and $|({}^\bullet j)| = 1, {}^\bullet({}^\bullet j) = \{p\}$, then:

$$\Gamma_j^{oac} = \Gamma_p^{oac} + DI(j)/ST(j)$$

$$\Gamma_j^{\tilde{o}ac} = \Gamma_p^{\tilde{o}ac} + DI(j)$$

$$\Gamma_j^{oa\tilde{c}} = \Gamma_p^{oa\tilde{c}} + DI(j)/ST(j)$$

$$\Gamma_j^{\tilde{o}a\tilde{c}} = \Gamma_p^{\tilde{o}a\tilde{c}} + DI(j)$$

(2) If $|({}^\bullet j)| = 1, |{}^\bullet({}^\bullet j)| > 1$, then:

$$\Gamma_j^{oac} = \max\{\Gamma_p^{oac} + DI(j)/ST(j) | p \in {}^\bullet({}^\bullet j)\}$$

$$\Gamma_j^{\tilde{o}ac} = \max\{\Gamma_p^{\tilde{o}ac} + DI(j) | p \in {}^\bullet({}^\bullet j)\}$$

$$\Gamma_j^{oa\tilde{c}} = \max\{\Gamma_p^{oa\tilde{c}} + DI(j)/ST(j) | p \in {}^\bullet({}^\bullet j)\}$$

$$\Gamma_j^{\tilde{o}a\tilde{c}} = \max\{\Gamma_p^{\tilde{o}a\tilde{c}} + DI(j) | p \in {}^\bullet({}^\bullet j)\}$$

(3) If $|({}^\bullet j)| > 1, |{}^\bullet({}^\bullet j)| > 1$, then:

$$\Gamma_j^{\tilde{o}ac} = \min\{\Gamma_p^{\tilde{o}ac} + DI(j) | p \in {}^\bullet({}^\bullet j)\}$$

$$\Gamma_j^{\tilde{o}a\tilde{c}} = \max\{\Gamma_p^{\tilde{o}a\tilde{c}} + DI(j) | p \in {}^\bullet({}^\bullet j)\}$$

$$\Gamma_j^{oac} = \min\{\Gamma_p^{oac} + DI(j)/ST(j) | p \in {}^\bullet({}^\bullet j)\}$$

$$\Gamma_j^{oa\tilde{c}} = \max\{\Gamma_p^{oa\tilde{c}} + DI(j)/ST(j) | p \in {}^\bullet({}^\bullet j)\}$$

(4) If $|(^\bullet j)| > 1, |^\bullet(^\bullet j)| > 1$, and $\forall y \in^\bullet, (y,) \in F''$, then:

$$\Gamma_j^{\tilde{o}\tilde{a}c} = \max\{\Gamma_p^{\tilde{o}\tilde{a}c} + DI(j)|p \in {}^\bullet({}^\bullet j)\}$$

$$\Gamma_j^{\tilde{o}\tilde{a}\tilde{c}} = \max\{\Gamma_p^{\tilde{o}\tilde{a}\tilde{c}} + DI(j)|p \in {}^\bullet({}^\bullet j)\}$$

$$\Gamma_j^{oac} = \min\{\Gamma_p^{oac} + DI(j)/ST(j)|p \in {}^\bullet({}^\bullet j)\}$$

$$\Gamma_j^{oa\tilde{c}} = \min\{\Gamma_p^{oa\tilde{c}} + DI(j)/ST(j)|p \in {}^\bullet({}^\bullet j)\}$$

In the formulas, notation "/" means "or". If p is an optional task, then $ST(p)$ is used, else, $DI(p)$ is used.

$\Gamma_p^{oac} - DI(p), \Gamma_p^{\tilde{o}\tilde{a}\tilde{c}} - DI(p), \Gamma_p^{oa\tilde{c}} - DI(p)/ST(p), \Gamma_p^{\tilde{o}\tilde{a}\tilde{c}} - DI(j)/ST(j)$ are used to compute the earliest time at which p start to execute according to the four basic paths. Using the computed time information, we can obtain the time performance of a business process and predict the implementation of real-time predictability, and judge whether the execution of a business process can meet the expectation. When the system is running, we can determine whether the current execution can meet the time expectation based on the known time information, and take the appropriate measures to avoid the occurrence time violation. For example, abandon the implementation of some optional tasks or pick faster alternative paths.

4 A Case Study

In this paper, a TBPN including all the structures in Sect. 2 is shown in Fig. 3. The execution delays of the tasks are shown in Table 1. The reachable marking graph is shown in Fig. 4. By analyzing the reachable marking graph, it is easy to judge that the TBPN is structural correct. By the formulas in Sect. 3, we can compute the earliest end and start time information of the tasks in the TBPN. The finished time information of the four paths can also be obtained. From Table 2, we know that the entire business process model can be finished at least 21 time units, and the slowest path consumes 72 time units. The earliest start time of each task is shown in Table 3. The business process must be executed strictly in accordance with the
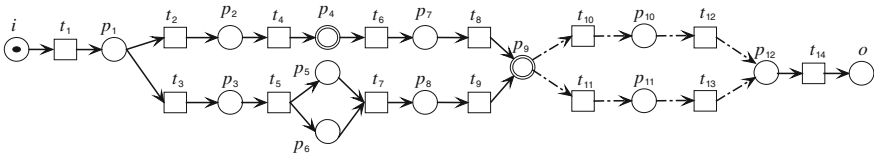
Fig. 3 An illustrated example

Table 1 Execution time of tasks

	i	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	o
DI	0	2	3	5	2	9	15	2	5	8	4	18	11	0
ST					0					0				

$Mi(1,0,0,0,0,0,0,0,0,0,0,0,0,0)$

t_1

$M1(0,1,0,0,0,0,0,0,0,0,0,0,0,0)$

$t_2 \qquad t_3$

$M2(0,0,1,0,0,0,0,0,0,0,0,0,0,0) \quad M3(0,0,0,1,0,0,0,0,0,0,0,0,0,0)$

$t_4 \qquad t_5$

$M4(0,0,0,0,1,0,0,0,0,0,0,0,0,0) \quad M5(0,0,0,0,0,1,1,0,0,0,0,0,0,0)$

$t_6 \qquad t_7$

$M6(0,0,0,0,0,0,0,1,0,0,0,0,0,0) \quad M7(0,0,0,0,0,0,0,0,1,0,0,0,0,0)$

$t_8 \qquad t_9$

$M8(0,0,0,0,0,0,0,0,0,1,0,0,0,0)$

$t_{10} \qquad t_{11}$

$M9(0,0,0,0,0,0,0,0,0,0,1,0,0,0) \quad M10(0,0,0,0,0,0,0,0,0,0,0,1,0,0)$

$t_{12} \qquad t_{13}$

$M11(0,0,0,0,0,0,0,0,0,0,0,0,1,0)$

t_{14}

$Mo(0,0,0,0,0,0,0,0,0,0,0,0,0,1)$
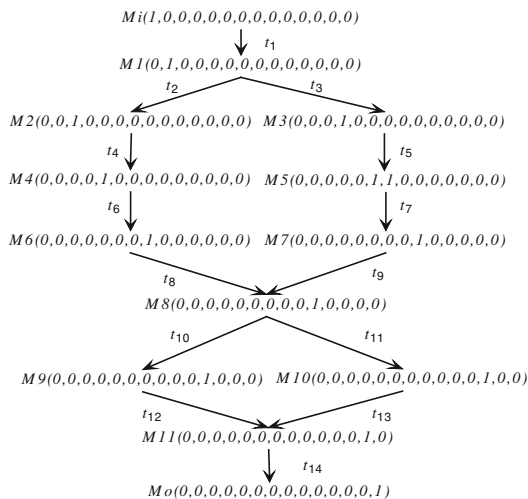
Fig. 4 The reachable marking graph

Table 2 The earliest end time of the tasks in the TBPN

	i	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	o
Γ_p^{oac}	0	2	5	7	5	16	22	7	27	7	11	25	22	22
$\Gamma_p^{\tilde{o}\tilde{a}c}$	0	2	5	7	7	16	22	9	27	17	21	29	40	40
$\Gamma_p^{oa\tilde{c}}$	0	2	5	7	5	16	22	7	27	27	31	45	42	42
$\Gamma_p^{\tilde{o}\tilde{a}\tilde{c}}$	0	2	5	7	7	16	22	9	27	35	39	53	64	64

time expectation. Once violate the time predictability, appropriate strategy must be taken. For example, if p_7 has just finished at time $\Gamma_{p_7}^{oa\tilde{c}} = \Gamma_{p_7}^{oac} < 8 < \Gamma_{p_7}^{\tilde{o}\tilde{a}c} = \Gamma_{p_7}^{\tilde{o}\tilde{a}\tilde{c}}$.

Then, we can judge that this business instance violate the *oac* and *oac̃* paths. Abandoning the implementation of some optional tasks or picking faster alternative paths must be done to meet the time predictability. When the system is running, given a deadline of a task or the entire business process, we can also dynamically determine whether the business processes meet the time expectation. The calculation process still follows the formulas in Sect. 3.

Table 3 The earliest start time of the tasks in the TBPN

	i	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	o
$\Gamma_p^{oac} - DI(p)/ST(p)$	0	0	2	2	5	7	7	5	22	7	7	7	11	11
$\Gamma_p^{\tilde{o}\tilde{a}c} - DI(p)$	0	0	2	2	5	7	7	7	22	9	17	17	29	41
$\Gamma_p^{oa\tilde{c}} - DI(p)/ST(p)$	0	0	2	2	5	7	7	5	22	27	27	27	31	35
$\Gamma_p^{\tilde{o}\tilde{a}\tilde{c}} - DI(j)$	0	0	2	2	5	7	7	7	22	27	35	35	53	62

5 Conclusions

To realize the reliable execution of business processes, it is need to verify the correctness and real-time predictability of business process models. This work's contribution is a formal model called Timed Business Process Net (TBPN) for modeling and analyzing business processes with time factors. The formal definitions are given. This method can provide a basic technology support to guarantee the real-time predictability of business processes. The proposed model can also be applied to other systems such as real-time control systems. The next work is to develop a prototype system based on TBPN.

References

1. Van der Aalst W, ter Hofstede A, Weske M (2003) Business process management: A survey. Bus Process Manage 2003:1019–1019
2. Weske M (2012) Business process management architectures. Bus Process Manage 2012:333–371
3. Guha S, Kettinger WJ, Teng JTC (1993) Business process reengineering. Inf Syst Manage 10:13–22
4. Chen H, Wang J, Dong W (2003) High confidence software engineering technologies. Acta Electronica Sinica 31:1933–1938
5. Arlbjorn JS, Haug A (2010) Business process optimization. Academica, California
6. Becker J, Kugeler M, Rosemann M (2003) Process management: a guide for the design of business processes. Springer, London
7. Hepp M, Roman D (2007) An ontology framework for semantic business process management. In: Proceedings of wirtschaftsinformatik 2007
8. Iwashita M, Asaki K, Yamamoto N (2007) Business process reconstruction method, and its program and computer. Google Patents
9. Paris S, Sillion FX, Quan L (2006) A surface reconstruction method using global graph cut optimization. Int J Comput Vision 66:141–161
10. Vergidis K, Tiwari A, Majeed B (2008) Business process analysis and optimization: beyond reengineering. Syst Man Cybern Part C: Appl Rev IEEE Trans 38:69–82

11. Eder J, Panagos E, Pozewaunig H, Rabinovich M (1999) Time management in workflow systems. BIS 99:265–280
12. Lanz A, Weber B, Reichert M (2010) Workflow time patterns for process-aware information systems. Enterp Bus Process Inf Syst Model 2010:94–107
13. Li JQ, Fan YS, Zhou MC (2003) Timing constraint workflow nets for workflow analysis. Syst Man Cybern Part A: Syst Humans IEEE Trans 33:179–193
14. Ling S, Schmidt H (2000) Time petri nets for workflow modelling and analysis. In: IEEE international conference on systems, man, and cybernetics, pp 3039–3044
15. Zhehui W (2006) Introduction to petri nets. Machine Press, Beijing

Automatic Text Classification Based on Hidden Markov Model and Support Vector Machine

Li Wang and Li Li

Abstract This paper researches the technology of text classification, and proposes a two-layer structure automatic text classification based on HMM and SVM. The given text is classified with HMM classifiers first to select the most likely two classes. Then these classes as SVM input are processed. Finally the given text is classified into the corresponding category with SVM classifier. The experimental results show that this method is more efficient for text classification in recognition ratio.

Keywords Hidden markov model (HMM) · SVM · Text classification · Term frequency-inverse document frequency (TF-IDF)

1 Introduction

At present, the amount of information is growing rapidly, and information behaves formal diversity and is out of order. Using text classification technology, enormous and lack of structural text data can be organized into a standard text data, to solve the problem of information chaos, and to help people to get information fast, accurately and comprehensively.

Text categorization is the process that classifies the given text into one or more predefined categories. In recent years, automatic text classification technology research has made great progress. Researchers put forward a lot of effective classification model, such as based on the kind of center method, Naive Bayes algorithm, SVM, KNN, neural network, decision tree, Boosting etc., [1–3].

L. Wang (✉) · L. Li
School of Computer Science and Engineering, Anhui University of Science
and Technology, Huainan, Anhui 232001, China
e-mail: liwang@aust.edu.cn

These models use occurrence frequency of the feature words as the basis of categorization algorithm, and don't take full account of the semantic information of the feature words in the text [4]. So in this paper, we combine the advantages of HMM and SVM, and present a two-layer structure text classification model based on HMM-SVM. The model considers the feature words' semantic relatedness, and can automatically classify text.

2 Text Representation and Feature Extraction

2.1 Text Representation

In the course of information processing, text representation mainly uses the vector space model (VSM).The model uses vector to represent the text: (W_1, W_2, …, W_n), there W_i is the i feature item weight [1–3]. In this model the term and term weights can be adjusted according to the specific application. So the model has flexibility and adaptability.

To express the given text as a vector in vector space, we carry out word segmentation first. After word segmentation stage, all words are included in the vector as dimensions of vector. But there are many feature items, and some are useless, so we need to reduce the dimension by feature extraction. Many sophisticated feature selection methods following have been proposed.

(1) TF (Term Frequency)

Term Frequency is the number of occurrences of the term in the entire corpus and is denoted tf (t, d). The higher the tf is, the higher the importance (weight) is for the doc. In this way, the words term frequency less than a threshold are deleted. Thereby the dimensionality of the feature space is reduced. But sometimes the low frequency words contain more information. We don't simply delete those words.

(2) DF (Document Frequency)

Document frequency refers to the number of documents containing the term in the entire data set. The classifier removes the features that its document frequency is particularly low or particularly high according to the pre-set threshold value.

(3) IG (information gain)

Information gain is the value to determine a relation between term t and class C by the function. The more information the term t brings the system, the more important the term t is. IG is defined in Eq. (1).

$$IG(T) = -\sum_{i=1}^{n} P(C_i \log_2 P(C_i) + P(t) \sum_{i=1}^{n} P(C_i|t) \log_2 P(C_i|t)$$
$$+P(\bar{t}) \sum_{i=1}^{n} P(C_i|\bar{t}) \log_2 P(C_i|\bar{t}) \tag{1}$$

where P(t) is the probability of term t occurrence, and $P(C_i)$ is the probability of the i-th category, $P(C_i|t)$ is the probability that the i-th category is taken when term t appears.

(4) MI (mutual information)

MI measures how much information the presence/absence of a term t contributes to making the correct classification decision on C. High probability of occurrence in category C_i, and the low probability of occurrence in other categories, term t will obtain higher mutual information and is very likely to be selected as the characteristics of the category of C_i. MI calculation method is shown in Eq. (2).

$$MI(t, C) = \sum_{i} P(C_i) \log \frac{P(t|C_i)}{P(t)} \tag{2}$$

(5) χ2-test (Chi square test)

χ2 applied to test the independence of occurrence of the term and occurrence of the class. A high value of χ2 indicates that the hypothesis of independence, which implies that expected and observed counts are similar, is incorrect. If the two events are dependent, then the occurrence of the term makes the occurrence of the class more likely (or less likely), so it should be helpful as a feature. χ2 is defined as shown in Eq. (3).

$$x^2(t, C_i) = \frac{N \times (AD - CB)^2}{(A + B) \times (B + D) \times (A + C) \times (C + D)} \tag{3}$$

where N is the total number of documents as before. A is the number of documents which belong to the class C_i and contain term t. B is the number of documents which don't belong to the class C_i and contain term t. C is the number of documents which belong to the class C_i and don't contain term t. D is the number of documents which don't belong to the class C_i and don't contain term t.

(6) ECE (expected cross entropy)

Expected cross entropy reflects the distance between the probability distribution of the categories and the probability distribution of the categories under the conditions of term t occurrences. ECE is defined as shown in Eq. (4).

$$ECE(t, C) = P(t) \sum_i P(C_i|t) \log \frac{P(C_i|t)}{P(C_i)} \tag{4}$$

(7) WET (the weight of evidence for text)

The weight of evidence for text measures the difference between the probabilities of the classes and the conditional probabilities of the classes when terms are given. If P(Ci|t) is big and the probability of the corresponding category appearance is small, it indicates that term t have major impact on text classification system. So the calculated value of the function is big, and term t can be selected as the feature. WET is defined as shown in Eq. (5).

$$WET(t, C) = P(t) \sum_i P(C_i) \log \frac{P(C_i|t)(1 - P(C_i))}{P(C_i)(1 - P(C_i|t))} \tag{5}$$

2.2 Feature Extraction

Correct feature selection is the premise of the classification algorithm working properly. When the classification algorithm is selected, the appropriate feature extraction can reduce the vector dimension, and simplify the calculation, and has a decisive impact on the classification accuracy.

The feature extraction is defined as a function of one or more measurements, each of which specifies some quantifiable property of a document, and is computed such that it quantifies some significant characteristics of a document. This paper compares several feature selection methods in text categorization, and adopts an improved term frequency and inverse document frequency methods shown in Eq. (6) [5].

Equation (6) takes into account that in practice the same words in different categories of articles should have different importance, and contain a different amount of information. So there may be a few words which has higher classification information for articles.

$$W(t, d) = \frac{tf(t, d) \times \log(\frac{N}{n_t} + 0.01) \times W(w_k, c_i)}{\sqrt{\sum_{t \in d} \left[tf(t, d) \times \log(\frac{N}{n_t} + 0.01) \times W(w_k, c_i) \right]^2}} \tag{6}$$

3 Text Classification Based on HMM/SVM

HMM focuses on the expression of class internal similarity, while SVM focuses on the difference between categories. Therefore, it is feasible to design a classifier which combines these two methods. The text classification based on HMM/SVM can be divided in the following stages in Fig. 1.

- Text preprocessing. The preprocessing is an important part of the automatic text classification systems. Its performance will seriously affect the capacity and speed of the systems. The processing involves filtering numbers, punctuation, spaces and etc., of the text, extracting characteristics and normalizing, and removing stop words. The output of this stage is a set of words.
- Feature extraction. Feature extraction is a process that chooses a subset of features from the original features so that the feature space is optimally reduced according to a certain criterion. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input. In this paper we adopt an improved TF-IDF method to extract feature.
- Classifier training. Training-Set is the source of knowledge of a classifier and affects the training result and the comprehensive ability of a classification system. The text classifier consists of both the training and the testing phases. In the training or the learning phase, the system is presented with inputs and known outputs. The system is supposed to fix its parameters or learn. So the two-layer model is built.
- Classification. After the features are extracted, the built two-layer classifier can work. In the testing phase, the system is given unknown inputs ands is asked to classify these inputs into classes. First, the HMM classifier as an initial classifier may allow documents to map in two classes. Then SVM classifier may allow documents to map to a single class.

3.1 Classifier Based on HMM

HMM is a finite state automata with double random process, Markov chains and stochastic processes. A hidden Markov models can be succinctly defined as $\lambda = (A, B, \pi)$. There $\pi = \{\pi1, \pi2,\ldots, \piM\}$ is the initial state probability vector, where $\pi i = P (q1 = si)$, $1 \leq i \leq M$ (M is the number of observable state). A is

Fig. 1 The processing of the text classification based on HMM/SVM
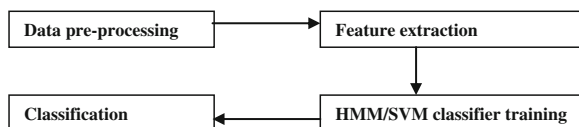
implied state transition probability matrix. B is the observed state transition probability matrix [6].

In this work, the training data set prepared in advance are used for HMM classifier training. Firstly we did pretreatment on text including word segmentation and removal of stop words in order to make the feature items more reasonable. And each document is converted into a set of words. Each word is weighted by the Eq. (6), and the weighted heavy words bigger than threshold setting are regarded as features. So the feature set is built up successfully.

The system needs to establish HMM models for training set. By means of the Baum—Welch algorithm,the initial state probabilities, state transition probabilities and observed the symbol probability are estimated, then HMM classifier for each class are establish.

Then using the forward algorithm, the output probability $P(O|\lambda i)$ of λi each observation sequence O in each HMM is calculated. The maximum output probability and the second larges output probability are selected as the input of SVM, and corresponding SVM classifier is working for further classification.

3.2 Classifier Based on SVM

SVM is a typical two types of problem identification. For multi-class, SVM and binary decision tree can be combined to form a multi-class classification [7, 8]. The multiple categories are divided into two classes of sets, and each set contains a plurality of categories. And the input is classified using two types of classifier progressively, to find out the optimal classification result eventually. Partial binary tree is generated based on the geometrical distribution of the training samples in the attribute space [8], as shown in Fig. 2.

Support vector machine(SVM) is to correctly classify samples into two parallel planes in input or feature space by optimal planes(lines), and the margin between the two classes is made to be the largest. In this work, SVM uses the improved inter-class similarity measurement d. Suppose there are two types of C_p and C_q, and number of samples are m. A Sample $X = [x_1, x_2, \ldots, x_m]^T$ belongs to the class C_p, where $xi = (a_{i1}, a_{i2}, \ldots, a_{in})$, $i = 1, 2, \ldots, m$. A Sample $Y = [y_1, y_2, \ldots, y_m]^T$

Fig. 2 Classification based on partial binary tree of SVM
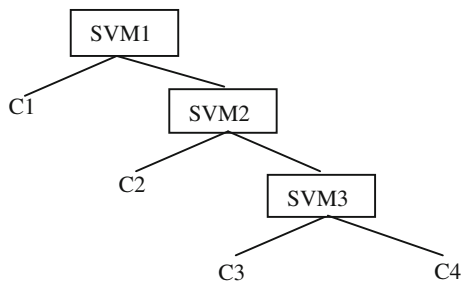
belongs to the class C_p, where $y_i = (b_{i1}, b_{i2},..., b_{in})$. Then $d(C_p, C_q)$ can be calculated by the Eqs (7), (8), (9).

$$d(C_p, C_q) = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} \frac{(a_{ij} - b_{ij})}{e(a_{ij}) + e(b_{ij})}} \tag{7}$$

$$e(a_{ij}) = \frac{|a_{ij}(C_p)|}{|a_{ij}|} \bullet \ln \frac{|a_{ij}(C_p)|}{|a_{ij}|} + \frac{|a_{ij}(C_q)|}{|a_{ij}|} \bullet \ln \frac{|a_{ij}(C_q)|}{|a_{ij}|} \tag{8}$$

$$e(b_{ij}) = \frac{|b_{ij}(C_p)|}{|b_{ij}|} \bullet \ln \frac{|b_{ij}(C_p)|}{|b_{ij}|} + \frac{|b_{ij}(C_q)|}{|b_{ij}|} \bullet \ln \frac{|b_{ij}(C_q)|}{|b_{ij}|} \tag{9}$$

In Eq. (8), $|a_{ij}|$ represents the occurrences of a_{ij} in X, and $|a_{ij}(C_p)|$ is the number the sample includes a_{ij} belongs to C_p, and rest of the variables are described Same as above. $e(a_{ij})$ denotes the entropy of a_{ij} for C_p, and $e(b_{ij})$ denotes the entropy of b_{ij} for C_q. The smaller the entropy value i.e., action level is, the greater the degree of action is.

SVM uses the RBF kernel function, as shown in Eq. (10).

$$K(x_i, x_j) = \exp(-\frac{\|x_i - x_j\|^2}{\sigma^2}) \tag{10}$$

4 Experimental Analysis and Conclusion

The experimental data are part of the documents selected from the 10 categories of Fudan University Chinese text corpus, and Table 1 shows document distribution of data. Words segmentation method is standard provided by Institute of Computational Linguistics of Peking University, and segmentation system of the Chinese Academy of Sciences is adopted. HMM is implemented by Matlab software, and SVM classifier is achieved using LIBSVM toolkit.

From Table 2, we may see that the average recognition rate is 76.71 %, and two level of recognition rate is 81.47 %. Through two-layer structure to classify text,

Table 1 The classes in the Fudan University Chinese text corpus with number of documents in training sets

Class	Number of documents	Class	Number of documents
Economy	519	Environment	327
Sports	359	Art	294
Computer	628	Outer- space	192
Politics	405	History	466
Agriculture	550	Military affairs	75

Table 2 A comparison between the results

Class	Recognition rate (%)		Recognition speed(word/second)	
	SVM	HMM/SVM	SVM	HMM/SVM
Economy	79.71	83.79	40.92	35.78
Sports	78.23	82.56	41.79	36.61
Computer	65.32	74.25	34.76	30.05
Politics	70.48	76.41	38.54	31.14
Agriculture	80.12	84.56	36.97	30.86
Environment	77.49	81.73	40.13	34.23
Art	81.23	85.04	41.82	35.91
Outer-space	78.94	81.91	40.47	33.94
History	76.58	80.92	39.24	33.03
Military affairs	79.04	83.54	38.98	34.26

the system has a certain improvement in recognition rate. But this two-layer-structure classic method has to spend long time for model for HMM and SVM. Later we focus on how to optimize the structure and shorten the time.

References

1. Liu Y-Z, Lin Y-P, Chen Z-P (2004) Text information extraction based on hidden markov model. J Syst Simul 16(3):507–509
2. Zhou S-X, Lin Y-P, Wang Y-N (2007) Text information extraction based on the second-order hidden markov model. ACTA Electronica Sinica 11:2226–2231
3. Luo S-H, Ouyang W-M (2007) HMM based text categorization. Comput Eng Appl 10:179–181
4. Yang J, Wang H-H (2010) Text classification algorithm based on hidden Markov model. J Comput Appl 9:2348–2351
5. Luo X, Xia D-L, Yan P-L (2005) Improved feature selection method and TF-IDF formula based on word frequency differentia. J Comput Appl 5(9):2031–2033
6. Sebastiani F (2002) Machine learning in automated text categorization. J ACM Comput Surv 34(1):1–47
7. Tsang IW, Kwok JT, Cheung P-M (2005) Core vector machines: fast SVM training on very large data sets. J Mach Learning Res 363–392
8. Gulinazi L, Sun T-L (2011) The application of multi-class SVM based binary tree in web text categorization. J XinJiang Univ (2):100–104

Existence of Periodic Solutions for n-Dimensional p-Laplacian Equation with Multiple Deviating Arguments

Jinbo Ni and Juan Gao

Abstract In this paper, we mainly discuss the existence of periodic solutions for n-Dimensional p-Laplacian differential equation with multiple deviating arguments. Under some broader sign conditions, new existence results are obtained by using the degree theory.

Keywords p-Laplacian equation · Delay equation · Periodic solution · Degree theory

1 Introduction

Second-order functional differential equations have wide applications in the system control and the existence of periodic solutions for this problem has been deeply concerned. The periodic solutions of the Lienard equation have been widespread concerned by many experts and scholars, and some important results are obtained with the deepening of research [1–4]. The authors of [5–8] studied the existence of periodic solutions for Lienard equation with a deviating argument as follows

$$x'' + f(x)x' + g(x(t - \tau(t))) = e(t)$$

The periodic solutions of the following p-Laplacian equations

J. Ni (✉) · J. Gao
Department of Mathematics, Anhui University of Science and Technology,
Huainan 232001 Anhui, People's Republic of China
e-mail: nijinbo2005@126.com

J. Gao
e-mail: gaojuan800@sina.com

$$\left(\phi_p(x')\right)' + \frac{d}{dt} gradF(x) + g(t, x) = p(t)$$

where considered in papers [9, 10], where $\phi_p(x) = |x|^{p-2}x$, $p \geq 2$. But few articles focus on the p-Laplacian equations with a deviating argument of the form

$$\left(\phi_p(x')\right)' + \frac{d}{dt} gradF(x) + g(x(t - \tau(t))) = p(t)$$

In general, systems with delay are more difficult than these without delay, so the research of the problems with delay has many limitations. Motivated by the above work, we shall study the p-Laplacian equations with multiple deviating arguments of the form

$$\left(\phi_p(x')\right)' + \frac{d}{dt} gradF(x) + g(t, x(t - \tau_1), x'(t - \tau_2)) = p(t) \qquad (1.1)$$

2 Preliminaries

For convenience, we introduce the following notations and lemmas.

Set $C_T^1 = \left\{x : x \in C^1(R, R), x(t + T) \equiv x(t)\right\}$.

Define the operator

$$L_p : domL_p \subset C_T^1 \rightarrow ([0, T], R), (L_px)(t) = \left(\phi_p(x'(t))\right)',$$

where $domL_p = \left\{x : x \in W^{2, p}([0, T], R^n), x(t + T) = x(t), t \in [0, T]\right\}$.

Define projection operator $Q : C_T \rightarrow C_T, (Qx)(t) = \frac{1}{T}\int_0^T x(s)ds$.

Lemma 1 [9] *Let $\Omega \subset C_T^1$ is an open and non-empty bounded set, $N : \overline{\Omega} \rightarrow C_T$ is L-compact operator. Assume that the following conditions are satisfied*

(A1) $L_px \neq \lambda Nx, \forall(x, \lambda) \in [domL_p \cap \partial\Omega] \times (0, 1)$;
(A2) $Nx \notin ImL_p, \forall x \in \ker L_p \cap \partial\Omega$;
(A3) $\deg_B\left(QN|_{\ker L_p}, \ker L_p \cap \partial\Omega, 0\right) \neq 0$.

Then the equation $L_px = Nx$ has at least one solution in $domL_P \cap \overline{\Omega}$.

3 Main Results

Theorem 1 (H_1) *There exists non-negative constants $\alpha, \beta, \gamma, m, r$ such that*

$$|g(t, x, y)| \leq g(t, x, y) + \alpha|x|^m + \beta|y|^r + \gamma, \quad \forall t \in [0, T], x, y \in R^n,$$

or

$$|g(t, \ x, \ y)| \leq \ - \ g(t, \ x, \ y) \ + \ \alpha|x|^m + \beta|y|^r + \gamma, \quad \forall t \in [0, T], x, y \in R^n,$$

(H_2) *There exist Lebesgue integrable function* $\mu_i^+(t)$, $\mu_i^-(t)$ *and constant* $M > 0$ *such that* $\forall i \in I_n$

$$\varepsilon_i g_i(t, \ x, \ y) > \mu_i^+, \ x_i > M, \quad \forall t \in [0, \ T], \quad y \in R^n,$$

$$\varepsilon_i g_i(t, \ x, \ y) < \mu_i^-, \ x_i < \ - \ M, \quad \forall t \in [0, T], \quad y \in R^n,$$

where $\varepsilon_i \in \{1, \ -1\}$, *and* $\mu_i^+(t)$, $\mu_i^-(t)$ *satisfy* $\int_0^T \mu_i^- \ (t)dt \leq 0 \leq \int_0^T \mu_i^+ \ (t)dt$;

(H3) *There exists non-negative constants* σ *such that* $\left|\frac{\partial^2 F(x)}{\partial x^2}\right| \leq \sigma$, $\forall x \in R^n$

If one of the following conditions is satisfied

(1) $r = m = p \ - \ 1$, $n^{\frac{p-1}{p}}\left(T\beta \ + \ \alpha n^{\frac{p-1}{p}}T^p\right) < 1$;

(2) $r = p \ - \ 1$, $m < p \ - \ 1$, $n^{\frac{p-1}{p}}T\beta < 1$;

(3) $r < p \ - \ 1$, $m \ = \ p \ - \ 1$, $\alpha n^{\frac{2p-2}{p}}T^p < 1$;

(4) $r < p \ - \ 1$, $m < p \ - \ 1$.

The Eq. (1.1) has at least one $T-$ period solution.

Define the operator $N : C_T^1 \ \rightarrow \ C_T$

$$N(x)(t) \ = \ - \ \frac{d}{dt} gradF(x) \ - \ g(t, x(t \ - \ \tau_1), x'(t \ - \ \tau_2)) \ + \ p(t), \ t \in [0, T]$$

Consider the homotopic equation of (1.1)

$$\left(\phi_p(x')\right)' = -\frac{d}{dt} gradF(x) - \lambda g(t, x(t \ - \ \tau_1), x'(t \ - \ \tau_2)) + \lambda p(t), \quad \lambda \in [0, 1]$$
$$(3.1)$$

obviously, the Eq. (3.1) is equivalent with the abstract equation

$$L_p x \ = \ \lambda N x, \ \lambda \in [0, \ 1].$$

Let $\Omega_1 \ = \ \left\{x \in domL_p : L_p x \ = \ \lambda N x, \ \lambda \in [0, \ 1]\right\}$, we obtain the following lemma.

Lemma 2 *If the conditions of theorem 1 are true, we have set* Ω_1 *is bounded*

Proof For $\forall x \in \Omega_1$, integrating (3.1) form 0 to T, we can get

$$\int_0^T g(t, \ x(t \ - \ \tau_1), \ x'(t \ - \ \tau_2))dt \ = \ 0. \tag{3.2}$$

and

$$\int_0^T g_i(t, x(t - \tau_1), x'(t - \tau_2))dt = 0, \quad \forall i \in I_n.$$

We claim that there exists $t_i \in [0, T]$ such that $|x_i (t_i)| \le M$. If not, there exists $j \in I_n$, such that $\min_{t \in [0, T]} |x_j(t)| > M$. Applying the condition (H2), we have

$$\int_0^T g_i(t, x(t - \tau_1), x'(t - \tau_2))dt \ne 0,$$

which contradicts with (3.2).

So

$$|x_i(t)| \le M + \int_0^T |x_i'(t)| dt \le M + T^{1 - \frac{1}{p}} \left(\int_0^T |x_i'(t)|^p dt \right)^{\frac{1}{p}}$$
$$\le M + T\|x'\|_0, \forall t \in [0, T], i \in I_n.$$

Further more

$$|x_i(t)| \le M + \int_0^T |x_i'(t)| dt \le M + T^{1 - \frac{1}{p}} \left(\int_0^T |x_i'(t)|^p dt \right)^{\frac{1}{p}}$$
$$\le M + T\|x'\|_0, \forall t \in [0, T], i \in I_n.$$

Hence

$$\|x\|_0 \le n^{\frac{1}{p}}(M + T\|x'\|_0). \tag{3.3}$$

Because of

$$\int_0^T \left| (\phi_p(x'))' \right| dt \le \int_0^T \left| \frac{\partial^2 F(x)}{\partial x^2} \right| |x'(t)| dt + \int_0^T |g(t, x(t - \tau_1), x'(t - \tau_2))| dt$$
$$+ \int_0^T |p(t)| dt \tag{3.4}$$

By hypotheses (H1), (H3), we immediately get

$$\int_0^T \left| (\varphi_p(x'))' \right| dt \le \sigma \int_0^T |x'(t)| dt + \alpha \int_0^T |x(t - \tau_1)|^m dt + \beta \int_0^T |x(t - \tau_2)|^r dt + (\gamma + \|p\|_0)T,$$
$$\le \sigma \|x'\|_0 T + \alpha \|x\|_0^m T + \beta \|x'\|_0^r T + (\gamma + \|p\|_0)T. \tag{3.5}$$

For $\forall i \in I_n$, there exists $\xi_i \in [0, T]$ such that $x_i'(\xi_i) = 0$.

Thus

$$
\begin{aligned}
|x_i'(t)|^{p-1} &= |\varphi_p(x_i'(t))| = \left| \int_{\xi_i}^t \left(\varphi_p(x_i'(t)) \right)' dt \right| \\
&\le \int_0^T \left| \left(\varphi_p(x_i'(t)) \right)' \right| dt \le \int_0^T \left| \left(\varphi_p(x_i'(t)) \right)' \right| dt
\end{aligned}
$$

$$
\begin{aligned}
|x'(t)| &= \left(\sum_{i=1}^n |x_i'(t)|^p \right)^{\frac{1}{p}} \le \left(\sum_{i=1}^n \left| \int_0^T \left| \left(\varphi_p(x'(t)) \right)' \right| dt \right|^{\frac{p}{p-1}} \right)^{\frac{1}{p}} \\
&\le n^{\frac{1}{p}} \left| \int_0^T \left| \left(\varphi_p(x'(t)) \right)' \right| dt \right|^{\frac{1}{p-1}}, \quad \forall t \in [0, T]
\end{aligned}
$$

Hence

$$
\|x'\|_0^{p-1} \le n^{\frac{p-1}{p}} \int_0^T \left| \left(\varphi_p(x'(t)) \right)' \right| dt \tag{3.6}
$$

By (3.3), (3.5), (3.6), we have

$$
\|x'\|_0^{p-1} \le n^{\frac{p-1}{p}} T \left[\sigma \|x'\|_0 + \beta \|x'\|_0^r + \alpha n^{\frac{m}{p}} \left(M + T \|x'\|_0 \right)^m \right] + n^{\frac{p-1}{p}} \left(\gamma + \|p\|_0 \right) T.
$$

Obviously, if any of (1)–(4) is true, there exists a constant $M_1 > 0$ irrelevant to λ such that $\|x'\|_0 \le M_1$. By (3.3), we can find a constant $M_2 > 0$ irrelevant to λ such that $\|x\|_0 \le M_2$. Thus, we know there is a constant $M_3 > 0$ irrelevant to λ such that $\|x\| = \max\{\|x\|_0, \|x'\|_0\} < M_3$, and the set Ω_1 is bounded.

Let $\Omega_2 = \{x \in \ker L_p : Nx \in \operatorname{Im} L_p\}$, we have the following conclusion.

Lemma 3 *If the conditions of theorem 1 are true, we know set Ω_2 is bounded.*

Proof For $\forall x \in \Omega_2$, we have $QNx = 0$, which implies

$$
\frac{1}{T} \int_0^T g(t, x(t - \tau_1), x'(t - \tau_2)) dt = \frac{1}{T} \int_0^T p(t)\, dt.
$$

So

$$
\int_0^T g_i(t, x(t - \tau_1), x'(t - \tau_2)) dt = 0, \quad \forall i \in I_n. \tag{3.7}
$$

Since $x \in \Omega_2$, $x \in \ker L_p = \{x : x = d, d \in R\}$, we get $\|x\| = \max\{\|x\|_0, \|x'\|_0\} = |d|$. Thus, we claim $|d_i| \le M (\forall i \in I_n)$.

Otherwise, there exists i_0, such that $|x_{i_0}| = |d_{i_i}| > M$, applying the condition (H2), $\int_0^T g_{i_0}(t, x(t - \tau_1), x'(t - \tau_2)) dt \ne 0$.

It is a contradiction with (3.7), by $|d| = \left(\sum_{i=1}^{n} |d_i|^2\right)^{\frac{1}{2}}$, we can get $|d|$ is bounded.

Then, Ω_2 is a bounded set.

For $\forall d \in R^n$, $x = d \in \ker L_p$, we know

$$\frac{1}{T}\int_0^T g(t, x(t - \tau_1), x'(t - \tau_2))|_{x=d}dt = \frac{1}{T}\int_0^T g(t, d, 0)dt, \quad i \in I_n, \text{ is true}$$

By (H2), we can obtain $\forall i \in I_n$

$$(-1)^{s_i} g_i(t, d, 0) > \mu_i^+(t), \quad d_i > M \tag{3.8}$$

$$(-1)^{s_i} g_i(t, d, 0) < \mu_i^-(t), \quad d_i < -M \quad \text{where } s_i \in \{0, 1\} \tag{3.9}$$

Let $\Omega_3 = \{x \in \ker L_p : -\lambda\Lambda x + (1 - \lambda)\}QNx = 0, \quad \lambda \in [0, 1]$,
where $\Lambda = diag\{(-1)^{s_1}, (-1)^{s_2}, \cdots, (-1)^{s_n}\}$, we have the following lemma.

Lemma 4 *If the conditions of theorem 1 are true, the set Ω_3 is bounded.*

Proof For $\forall x = d = (d_1, d_2, \cdots, d_n) \in \Omega_3$, we claim $|d| \le M(\forall i \in I_n)$. If not, there is a $\bar{x} = \bar{d} = (\overline{d_1}, \overline{d_2}, \cdots, \overline{d_n}) \in \Omega_3$, exists $i_0 \in I_n$, satisfied $\left|\overline{d_{i_0}}\right| > M$, by $\bar{x} = \bar{d} \in \Omega_3$, we have $-\bar{\lambda}\Lambda(\bar{d}) + (1 - \bar{\lambda})QN(\bar{d}) = 0, \quad \bar{\lambda} \in [0, 1]$,

On the other hand,

$$QN(\bar{d}) = \frac{1}{T}\int_0^T \left[p(t) - \frac{d}{dt}gradF(x)\Big|_{x=\bar{d}} - g(t, x(t-\tau_1), x'(t-\tau_2))\Big|_{x=\bar{d}}\right]dt$$

$$= -\frac{1}{T}\int_0^T g(t, \bar{d}, 0)dt,$$

So

$$(-1)^{s_i}\bar{\lambda}\,\overline{d_i} + \left(1 - \bar{\lambda}\right)\frac{1}{T}\int_0^T g_i(t, d, 0)dt = 0, \quad i \in I_n$$

Especially

$$(-1)^{s_{i_0}}\bar{\lambda}\,\overline{d_{i_0}} + (1 - \bar{\lambda})\frac{1}{T}\int_0^T g_{i_0}(t, \bar{d}, 0)dt = 0.$$

According to the above formula, we can get $\bar{\lambda} \ne 1$. Otherwise $\overline{d_{i_0}} = 0$ is contradicts with $\left|\overline{d_{i_0}}\right| > M$. By (3.8), (3.9) and $\left|\overline{d_{i_0}}\right| > M$, the left of the above formula may not be equal to 0, It is a contradiction. Therefore, for $\forall x = d \in \Omega_3$, $|d_i| \le M, \forall i \in I_n$ then, $|d| = \left(\sum_{i=1}^{n} |d_i|^2\right)^{\frac{1}{2}} \le \sqrt{n}M$. Thus, Ω_3 is a bounded set.

Proof of Theorem 1 *Suppose $\Omega \subset Y$ is a open and bounded set and $\cup_{i=1}^{3} \overline{\Omega_i} \subset \Omega$. Operator N is continuous and map bounded set to bounded set, then N is L-compact in Ω.*

Applying Lemma 2 and 3

(A1) $L_p x \neq \lambda N x, \quad \forall (x, \lambda) \in [domL_p \cap \partial\Omega] \times [0, 1]$;

(A2) $Nx \notin \mathrm{Im} L_p, \quad \forall x \in \ker L_p \cap \partial\Omega$, are true.

On the other hand, we get $H(x, \lambda) = -\lambda \Lambda x + (1 - \lambda) QNx$.

By Lemma 4, we have $H(x, \lambda) \neq 0, \forall x \in \ker L_p \cap \partial\Omega, \quad \lambda \in [0, 1]$

By the homotopy invariance property, we get that

$$\begin{aligned} \deg\left(QN\big|_{\ker L_p}, \ker L_p \cap \Omega, 0\right) &= \deg\big(H(\cdot, 0), \ker L_p \cap \Omega, 0\big) \\ &= \deg\big(H(\cdot, 1), \ker L_p \cap \Omega, 0\big) \\ &= \deg\big(\pm I, \ker L_p \cap \Omega, 0\big) \neq 0. \end{aligned}$$

So the condition (A3) of Lemma 1 is satisfied. Applying Lemma 1, we reach the conclusion. This completed the proof of Theorem 1.

References

1. Mawhin J (1972) An extension of a theorem of AC Lazer on forced nonlinear oscillations. J Math Anal Appl 40:20–29
2. Ge W (1991) On the existence of harmonic solution of the Lienard system. J Differ Equ 16(2):183–190
3. Peng S, Zhu S (2003) Existence and uniqueness of periodic solutions for periodically perturbed non-conservitive systems. Chin Ann Math 24(3):293–298
4. Zhang M (1999) Periodic solutions of damped differential systems with repulsive singular forces. Proc Amer Math Soc 127(2):402–407
5. Huang X (1999) On the existence of harmonic solution for the n-dimensional Lienard equation with delay. J Syst Sci Math Sci 19(3):328–335
6. Liu B, Yu J (2002) On the existence of harmonic solution for the n-dimensional Lienard equation with delay. Acta Math Sci 22A(3):323–331
7. Lu S, Ge W (2004) Periodic solutions for a kind of lienard equation with a deviating argument. J Math Anal Appl 289:231–243
8. Li Y (1998) 2π periodic solutions of delay system of Duffing type. Pure Appl Math 14(2):23–27
9. Manasevich R, Mawhin J (1998) Periodic solutions for non-linear systems with p-Laplacian operators. J Differ Equ 145:367–393
10. Liu B, Yu J (2003) On the existence of solution for the periodic boundary value problems with p-Laplacian operator. J Syst Sci Math Sci 23(1):76–85

Approximate Completed Trace Equivalence of Linear Algebra Transition Systems

Hao Yang, Anping He, Zhiwei Zhang, Shihan Yang and Yang liu

Abstract We proposed the approximate completed trace equivalence of Linear Algebra Transition Systems(LATS); it is inefficient to eliminate states. We use linear polynomial programs to describe system actions, and then the completed trace equivalence of inhomogeneous linear transition systems is established. Subsequently, it proposed the theory and algorithm of approximate completed trace equivalence of linear algebra transition systems.

Keywords Linear algebra transition systems · Approximate completed trace equivalence · Linear polynomial programs

1 Introduction

The definition of algebraic transition system is proposed by Sriram Sankaranarayanan, which is specialized by general transition system, so algebraic transition system is a class of transition system. Transition system [1] is a basic concept in computer system model, which is used to study the real-time systems, hybrid systems. This paper is deeper research on our job before in Ref [2, 3].

Ref. [4] proposed a congruence format for completed traces. Ref. [5] studied Basic Process Algebra with interrupt modulo complete trace equivalence. Ref. [6] proposed complete trace equivalence can be axiomatized by the set of special axioms. Ref. [7] studied safe completed trace equivalence. They use symbolic

H. Yang · A. He (✉) · Z. Zhang · S. Yang
Guangxi Key Laboratory of Hybrid Computation and IC Design Analysis, Guangxi
University for Nationalities, Nanning, China
e-mail: hapetisat@gmail.com

H. Yang · Z. Zhang · Y. liu
Chengdu Institute of Computer Application Chinese Academy of Sciences, Chengdu, China

reasoning method to study completed trace equivalence. In this paper, numerical computation method is used to study completed trace equivalence on linear algebra transition systems.

2 Preliminaries

In this paper, we proposed the linear algebra transition systems(LATS). It uses linear polynomial programs to replace labels, in addition to overcome the shortcoming of the labeled transition system.

Let R be the set of reels, $x_i(i = 1,\ldots,n) \in R$ and $\dot{x}_i(i = 1,\ldots,n) \in R$ be the variables, $a_{ij}(i = 1,\ldots,n; j = 1,\ldots,n) \in R$ and b_i $(i = 1,\ldots,n) \in R$ be the coefficients. A linear polynomial program

$$\begin{cases} \dot{x}_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + b_1 \\ \dot{x}_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + b_2 \\ \qquad\qquad \cdots \\ \dot{x}_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + b_n \end{cases}$$

can be abbreviated as $\dot{X} = AX + b = F(X)$, where $\dot{X} = (\dot{x}_1, \dot{x}_2, \ldots, \dot{x}_n)^T$, $X = (x_1, x_2, \ldots, x_n)T$ and $b = (b_1, b_2, \ldots, b_n)^T$,

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}$$

Lemma 1 *In Ref.14, we got A non-homogeneous linear algebra transition system can be transited to a homogeneous linear algebra transition system.*

The paper rest we only talk about homogeneous linear algebra transition system.

3 Completed Trace Equivalence of Linear Algebra Transition Systems

A linear algebra transition systems, from state q_1 to state q_2 by executes action s, it can be written as (q_1, s, q_2) or $q_1 \xrightarrow{\text{S}} q_2$. Successor function succ expresses transition system successor state, $q_2 = succ(q_1, s)$. A sequence of executions $q_1 \xrightarrow{s_1} q_2 \xrightarrow{s_2} q_3 \ldots \xrightarrow{s_n} q_n + 1 \ (n > 0)$ of a transition system is composed by a sequence of states $\overline{q} = q_1 q_2 \ldots q_{n+1}$ and a sequence of actions $\overline{s} = s_1 s_2 \ldots s_n$. q_{n+1}

can be written as $q_{n+1} = succ\ (q_1, \bar{s})$. $\bar{s}$ is a trace of the transition system. $|\overline{s}|$ is a set of all actions in the sequence of executions $\bar{s}$, $|\overline{s}| = \{s_1, s_2, \ldots, s_n\}$. Q_1 is a set of the trace reached states, $Q_1 = \{q_1, q_2, \ldots, q_{n+1}\}$. Two transition systems are trace equivalent if they have the same traces. A completed trace of a transition system is a sequence of actions such that $q_1 \xrightarrow{s_1} q_2 \xrightarrow{s_2} q_3 \ldots \xrightarrow{s_n} q_n + 1\ (n > 0)$ and q_{n+1} cannot execute any action. Two transition systems are completed trace equivalence if they are trace equivalent and have the same completed traces. Symbol $= CT$ represents completed trace equivalence.

The completed trace equivalence of two labeled transition systems

The completed trace equivalence of labeled transition systems can only change the set of states and it cannot change the set of labels.

In the Linear Algebra Transition Systems, linear polynomial programs are used to describe real trace $\bar{s} = s_1 s_2 \ldots s_n (n > 0)$ and approximate trace $\bar{t} = t_1 t_2 \ldots t_m (m > 0)$. Let s_i: $\dot{X} = A(i)X + b(i) = F_i\ (X)(1 \le i \le n)$ and t_i: $\dot{X} = C(i)X + d(i) = G_i\ (X)(1 \le i \le m)$ be the linear polynomial programs, then $\bar{s} : \dot{X} = F_n\ (F_{n-1}\ (\ldots F_1\ (X)))$ and $\bar{t} : \dot{X} = G_m\ (G_{m-1}\ (\ldots G_1\ (X)))$. The trace $\bar{s} = s_1 s_2 \ldots s_n$ and the trace $\bar{t} = t_1 t_2 \ldots t_m$ are equivalent if $\bar{s} : \dot{X} = F_n\ (F_{n-1}\ (\ldots F_1\ (X)))$ and $\bar{t} : \dot{X} = Gm\ (Gm - 1(\ldots G1\ (X)))$ are equivalent.

4 Theory and Algorithm of Approximate Completed Trace Equivalence

The online version of the volume will be available in LNCS Online. Theorem 1 If $x_1, x_2, \ldots, x_n$ are three non-negative numbers, $n \ge 1$, then $\frac{x_1 + x_2 + \cdots + x_n}{n} \ge \sqrt[n]{x_1 x_2 \ldots x_n}$. The equal sign is tenable if all xi are equal.

Theorem 2 *If* $x_1, x_2, \ldots, x_n$ *are three non-negative numbers,* $n > 1$, $1 > x_i > 0 (i = 1, \ldots, n)$, *then*

$$(1 + x_1)\ (1 + x_2) \ldots (1 + x_n) \le \left(1 + \frac{x_1 + x_2 + \cdots x_n}{n}\right)^n < e^{x_1 + x_2 + \cdots + x_n},$$

where e *is the natural logarithm bottom* $\ln x = \log_e x$.

In the actual measurement, linear polynomial programs always have inevitable errors. In this paper, the approximate research of real traces is based on the initial state q_0 of Linear Algebra Transition Systems

Let q_0 be the initial state of a Linear Algebra Transition Systems. Let X_0 be the initial state value. A sequence of real executions of a Linear Algebra Transition Systems is $q_0 \xrightarrow{s_1} q_1 \xrightarrow{s_2} q_2 \ldots \xrightarrow{s_n} q_n\ (n > 0)$. The real trace is $\bar{s} = s_1 s_2 \ldots s_n$, where $s_i : \dot{X} = (A(i) + \delta A(i)\ )X + (b(i)\ + \delta b(i)) = F_i(X)\ (1 \le i \le n)$. The real value of state q_i is

$Xi + \delta Xi = Fi\,(Fi - 1(\ldots F1\,(X))) = (A(1) + \delta A(1)\,)\,(A(2) + \delta A(2)\,)\ldots(A(i) + \delta A(i)\,)X0$
$\qquad + (A(1) + \delta A(1)\,)\,(A(2) + \delta A(2)\,)\ldots(A(i - 1) + \delta A(I - 1)\,)\,(b(i) + \delta b(i))$
$\qquad + \ (A(1) + \delta A(1)\,)\,(A(2) + \delta A(2)\,)\ldots(A(i - 2) + \delta A(i - 2)\,)\,(b(i - 1)$
$\qquad + \delta b(i - 1)) + \cdots + (b(1) + \delta b(1)\,).$

A sequence of approximate executions of a LATS is $q_0 \xrightarrow{t_1} q_1 \xrightarrow{t_2} q_2 \ldots \xrightarrow{t_n} q_n\,(n > 0)$. The approximate trace is $\bar{t} : t_1 t_2 \ldots t_n$, where $t_i : \dot{X} = A(i)X + b(i) = G_i(X)\,(1 \leq i \leq n)$. The approximate value of state q_i is $X_i = G_i\,(G_{i-1}\,(\ldots G_1\,(X))) = A(1)\,A(2)\ldots A_{(i)}\,X0 + A(1)\,A(2)\ldots A(i - 1)b(i) + A(1)\,A(2)\ldots A(i - 2)\,b(i - 1) + \ldots + b(1)$. Let $\|\delta A(i)\| = h_i \|A(i)\|$, $\|\delta b(i)\| = g_i \|b(i)\|, 1 > \ h_i, \ g_i > 0 (i = 1, \ldots, n)$. Then

$$\| \delta X_i \| = \| (X_i + \delta X_i) - X_i \| \leq [(1 + h_1)\ldots(1 + h_i) - 1]$$
$$\| A(1) \| \ldots \| A(i) \| \| X0 \| + [(1 + h_1)\ldots(1 + h_{i-1})(1 + g_i) - 1]$$
$$\| A(1) \| \ldots \| A(i - 1) \| \| b(i) \| + [(1 + h_1)\ldots(1 + h_{i-2})(1 + g_{i-1}) - 1]$$
$$\| A(1) \| \ldots \| A(i - 2) \| \| b(i - 1) \| + \cdots + (1 + h_1) \| b(1) \| < e^{h_1 + \cdots + h_i} - 1)$$
$$\| A(1) \| \ldots \| A(i) \| \| X0 \| + (e^{h_1 + \cdots + h_{i-1} + g_i} - 1) \| A(1) \| \ldots \| A(i - 1) \| \| b(i) \| + (e^{h_1 + \cdots + h_{i-2} + g_{i-1}} - 1)$$
$$\| A(1) \| \ldots \| A(i - 2) \| \| b(i - 1) \| + \cdots + (e^{h_1} - 1) \| b(1) \| = W_i.$$

$\frac{W_i}{\|X_i\|}$ is the enlargement relative error of state q_i's approximate value.

Theorem 3 *Let ε be the error allowed value. When $n = 1, i = 1$, the real trace $\bar{s} = $ s1 and approximate trace $\bar{t} = $ t_1 are approximate equal if $\frac{W_1}{\|X_1\|} < \varepsilon$. When $n > 2, \ n > i \geq 1$, the real trace $\bar{s} = $ s1s2\ldots sn and approximate trace $\bar{t} = t_1 t_2 \ldots t_n$ are approximate equal, if all real traces $s_1 s_2 \ldots s_n$ and approximate traces $t_1 t_2 \ldots t_n$ are approximate equal and $\frac{W_n}{\|X_n\|} < \varepsilon$.*

The approximate of traces can reduce data storage bits and the number of inhomogeneous linear polynomial programs. Two different real linear polynomial programs can regard as the same approximate linear polynomial program by the approximate of traces. The approximate of traces can reduce data storage bits while the equivalence of traces cannot do it.

Definition 3 Two LATS s are approximate equal if their traces are all approximate equal. One LATS is called approximate system of another LATS.

Definition 4 If the approximate system of LATS and another LATS are equivalent, two LATS are approximate completed trace equivalent. The approximate completed trace is expressed by the symbol $\approx CT$.

5 Conclusion

In this paper, the theory and algorithm of approximate completed trace equivalence of LATS are proposed. Under certain conditions, the approximate completed trace equivalence of LATS is more effective on eliminating states than the completed trace equivalence of labeled transition systems.

References

1. Sun Y, Liu F, Mei S (2010) Polynomial approximate of a nonlinear system and its application topower system (I): theoretical justification. Electric Mach Control Harbin Univ Sci Technol 14(8):19–23
2. Yang H, Wu J, Zhang Z (2012) Approximate completed trace equivalence of inhomogeneous linear transition systems. IJACT 4(8):58–66
3. Zhang Z, Wu J-Z, Yang H Approximation to Linear Algebra Transition system. Communications in computer and information science, Vol 304. Emerging Intelligent Computing Technology and Applications, Part 10, pp 336–343
4. Klin B (2004) From bialgebraic semantics to congruence formats. In: Proceedings of the workshop on structural operational semantics, pp 3–37
5. Aceto L, Capobianco S, Ingólfsdóttir A (2007) On the existence of a finite base for complete trace equivalence over BPA with interrupt [Online]. Available: http://www.brics.dk/RS/07/5/BRICS-RS-07-5.pdf, February 2007
6. de Frutos-Escrig D, Gregorio-Rodríguez C, Palomino M (2008) Coinductive characterisations reveal nice relations between preorders and equivalences. Electron Notes Theoretical Comput Sci 212:149–162
7. Alvim M, Andrés M, Palamidessi C, van Rossum P (2010) Safe equivalences for security properties. Int Fed Inf Process 2010:55–70

A Soft Subspace Clustering Algorithm Based on Multi-Objective Optimization and Reliability Measure

Zhisheng Bi, Jiahai Wang and Jian Yin

Abstract Subspace clustering finds clusters in subspaces of the data instead of the entire data space to deal with high-dimensional data. Most existing subspace clustering algorithms lean on just one single objective function. Single objective function is often biased. On the other hand, most existing subspace clustering algorithms are based on wrapper approach, which brings a negative effect on the quality of subspace clustering. This paper presents a soft subspace clustering algorithm based on multi-objective evolutionary algorithm and reliability measure, called R-MOSSC. Comparing with optimization of a scalar function combining multiple objectives, it does not need to determine weight hyperparameters, and offers a deep insight into the problem by obtaining a set of solutions. Further, reliability-based dimension weight matrix from filter approach is used to enhance the performance of subspace clustering. Simulation results show that R-MOSSC is better than existing algorithms.

Keywords Subspace clustering · Multi-objective optimization · Reliability measure

1 Introduction

Clustering analysis is a data analysis technique with numerous applications in machine learning, statistics, and pattern recognition fields in the past decades. These applications include gene expression analysis, metabolic screening, customer recommendation systems and text analysis, Kriegel et al. [1]. Clustering analysis aims to group data points into several clusters by their similarity so that

Z. Bi · J. Wang (✉) · J. Yin
Department of Computer Science, Sun Yat-sen University, Guangzhou,
People's Republic of China
e-mail: wjiahai@hotmail.com

data points in the same cluster are more similar to each other and dissimilar to the points in other clusters. In high-dimensional sparse data space, the difference of distances between two data points is small due to "curse of dimensionality". Therefore, conventional clustering algorithms fail in high dimensions [2]. A solution for this problem is subspace clustering, which finds clusters in subspaces of the data instead of the entire data space.

According to the weighting methods, subspace clustering algorithms can be grouped into two basic types: hard subspace clustering and soft subspace clustering [3]. In hard subspace clustering, equal non-zero weights are given to relevant dimensions and zero weights to irrelevant dimensions, while, in soft subspace clustering, features have different weighting values based on their importance for a particular cluster. More details can be found in [4]. This paper focuses on soft subspace clustering. Entropy weighting k-means (EWKM) [3] is one of the most classical soft subspace clustering algorithms. In EWKM, a k-means (KM) [5] like objective weighting function is proposed by employing both the maximum information entropy theory and within-cluster information. It minimizes the objective function by the KM clustering process with an additional step of computing the weight values. The computational complexity of EKWM is low, and it can converge to local optima in a finite number of iterations. Enhance soft subspace clustering (ESSC) [4] is another KM type subspace clustering algorithm. It combines within-cluster information and between-class information to improve the performance of subspace clustering. Lu et al. [6] pointed out that the performance of soft subspace clustering largely depends on the objective function and the search strategy. They proposed a particle swarm optimizer for variable weighting (PSOVW) to avoid the shortage of KM like subspace algorithms.

Clustering analysis aims to group data points into several clusters by their similarity. This is a loose concept. Hence, there are many different criterions to evaluate the clustering quality. Most existing subspace clustering algorithms lean on just one single objective function. However, single objective function is often biased. Therefore, evaluating subspace clustering quality by multiple objectives is necessary. Combining multiple objectives into a scalar objective function is a way for taking into account more than one objective, which is widely used in subspace clustering, such as EWKM and ESSC. A weakness of this method is that it is difficult to set weight hyperparameters. Besides, the trade-off between objectives is fixed by predefined weight hyperparameters. Thus, only one solution is returned by the algorithms optimizing a scalar objective function. Comparing with optimization of a scalar function, multi-objective optimization does not need to determine weight hyperparameters. Further, a set of solutions obtained by multi-objective optimization offers a deep insight into the problem.

On the other hand, most existing subspace clustering algorithms are based on wrapper approach [7]. They update the dimension weight matrix according to the centroids matrix which would be changed immediately. Thus, the likelihood of weight accuracy may not be maintained, which brings a negative effect on the quality of subspace clustering [7]. To overcome this shortage, Boongoen et al. [7] proposed filter-based subspace clustering algorithms by extending data reliability

measure. In this filter approach, the weight configuration updates according to the object-dimension association matrix, which is a collection of information about the similarity of an object (data point) and its neighbors. This information, called object-specific reliability, is the true characteristics of the local relevance extracted from the data set. It helps keep up the accuracy of dimension weight.

This paper presents a soft subspace clustering algorithm based on multi-objective evolutionary algorithm (MOEA). Further, reliability-based dimension weight matrix from filter approach [7] is used to enhance the performance of subspace clustering. Therefore, the proposed algorithm, called R-MOSSC, is a filter framework based on MOEA. In R-MOSSC, the object-specific reliability is extracted by only one scan of the data set in the beginning of the algorithm. Thus, the computational cost of updating dimension weight would much less than that in the wrapper approach. Then, a MOEA proposed recently, called archive-based micro genetic algorithm 2 (AMGA2) [8], is adopted to optimize two objective functions. The simulation results show that R-MOSSC is significantly better than the previously existing algorithms.

2 R-MOSSC

Let $\mathbf{U} = \left[u_{ij}\right]_{C \times N}$ be the partition matrix, $\mathbf{V} = \left[v_{ij}\right]_{C \times D}$ be the centroids matrix, and $\mathbf{W} = \left[w_{ij}\right]_{C \times D}$ be the dimension weight matrix, where C, N, D are the number of clusters, the number of data points and the number of dimensions, respectively.

2.1 Reliability-Based Dimension Weight Matrix

Let N_{jk}^{α} be the set of α neighbors in the kth dimension of the jth data point, D_{jk}^{α} be the average distance between the jth point and its α neighbors in the kth dimension. The object-dimension association matrix $\mathbf{AS}$ is defined as follows [7]:

$$D_{jk}^{\alpha} = \frac{1}{\alpha} \sum_{q \in N_{jk}^{\alpha}} \sqrt{(x_{jk} - q_k)^2}, \ AS_{jk}^{\alpha} = 1 - \left(D_{jk}^{\alpha} \Big/ \max_{\forall j,k} D_{jk}^{\alpha}\right). \tag{1}$$

In the initial stage, a set of C data points is selected randomly as initialize centroids of an individual. The dimension weight is computed according to the selected points directly:

$$w_{ik} = AS_{jk}^{\alpha} \Big/ \sum_{k'=1}^{D} AS_{jk'}^{\alpha} \tag{2}$$

In the iterations, the dimension weight is updated as follows:

$$w_{ik} = MS^{\alpha}_{jk} \Big/ \sum\nolimits_{k'=1}^{D} MS^{\alpha}_{jk'}, \; MS^{\alpha}_{jk} = \min_{q \in Cluster_i} AS^{\alpha}_{qk} \tag{3}$$

2.2 Multi-Objective Optimization for Subspace Clustering

A multi-objective minimization problem could be stated as follows:

$$\begin{aligned} \min F(x) F(x) &= \{f_1(x), f_2(x), \dots, f_m(x)\} \\ \text{s.t. } g_i(x) \le 0, \quad i = 1, 2, \dots, P \quad h_j(x) &= 0, \quad j = 1, 2, \dots, Q. \end{aligned} \tag{4}$$

where F is a set of objectives, x is a solution of this problem. To compare two solutions with multiple objectives, Pareto domination is defined as follows: A solution x is said to dominate solution y if for all $i = 1, 2, \dots, m, f_i(x) \le f_i(y)$ and there exists at least one $j \in \{1, 2, \dots, m\}$ that $f_j(x) < f_j(y)$. A solution x is called Pareto-optimal if it is not dominated by any other solutions. The set including all Pareto-optimal solutions is called Pareto set (PS). The image of the PS on the objective space is called Pareto front (PF).

Lots of multi-objective optimization algorithms have been proposed for multi-objective optimization, and details can be found in recent review [9]. AMGA2 [8] is a MOEA proposed recently. It works with a very small population extracted from a large archive. Thus, the number of function evaluations needed at every generation is reduced. On the other hand, it combines several novel optimization concepts reported in the literature recently to improve its search capability. AMGA2 is shown to be an excellent MOEA, which can generate a high-quality solution with limited computational resources [8]. Therefore, AMGA2 is adopted to search clusters in subspace by optimizing two objectives in this paper.

Objective functions. A number of objective functions or clustering validity indexes could be used to evaluate the quality of subspace clustering. The within-class compactness and the between-cluster separation are often considerate as objective functions in exiting clustering algorithms [3, 4]. In this paper, two objective functions based on the within-class compactness and the between-cluster separation are also adopted to define the clustering quality as follows:

$$f_1 = \sum_{i=1}^{C} \sum_{k=1}^{D} w_{ik} \sum_{j=1}^{N} u_{ij} \left(x_{jk} - v_{jk}\right)^2, \tag{5}$$

$$f_2 = 1 \Big/ \sum_{i=1}^{C} \left(\sum_{j=1}^{N} u_{ij}\right) \sum_{k=1}^{D} w_{ik} (v_{ik} - v_{0k})^2, \; v_{0k} = \frac{1}{N} \sum_{j=1}^{N} x_{jk}. \tag{6}$$

The first objective function Eq. (5) is the sum of weighting distance between data points and their centroids. It reflects the within-class compactness. The

smaller the value is, the higher similarity within a cluster is. The denominator of the second objective function Eq. (6) is the weighting between-cluster scatter. The larger the value is, the higher separation between clusters is. As a minimization problem, we defined f_2 with the reciprocal of it. Therefore, by minimizing these objectives simultaneously, points in the same cluster will become more similar to each other, and become more dissimilar to the points in other clusters.

Chromosome representation and population initialization. Let $P_g = \{x_{1,g}, x_{2,g}, \ldots, x_{m,g}\}$ be the archive of the gth generation. $x_{i,g} = (x_{i,g,1}, \ldots, x_{i,g,D}, x_{i,g,D+1}, \ldots, x_{i,g,C \times D})$ is ith individual in $\mathbf{P}_g$ represents a centroids matrix $\mathbf{V}$, the first D gens represent the coordinate of the first centroid of this individual, the second D gens represent the second centroid, and so on. Besides, R-MOSSC maintains the dimension weight matrix and the partition matrix for each individual.

In the initial stage, a set of C data points is selected randomly as initialize centroids of an individual. Then, the dimension weight of each cluster is calculated according to Eq. (2). After that, the assignment of each point is decided as follows:

$$u_{ij} = \begin{cases} 1, & if \ i = \arg \min_{q=1,\ldots,C} \sum_{k=1}^{D} w_{qk} (x_{jk} - v_{qk})^2. \\ 0, & otherwise. \end{cases} \tag{7}$$

At last, the objectives of each individual are evaluated by Eqs. (5)–(6).

Evolution operators. AMGA2 manages a set of solutions and updates them iteratively by three steps: create parent population, create offspring population and update the archive. First, a small set of individuals in the archive are selected to create the parent population by sorting with their rank and crowding distance. Next, new offspring solutions are generated with differential evolution operator and polynomial mutation operator. Then, the partition matrix $\mathbf{U}$, the weight matrix $\mathbf{W}$ and the centroids matrix $\mathbf{V}$ of new offspring solutions are updated one after another by one iteration of R-KM [7]. After that, the new offspring solutions are evaluated and added to the archive. Finally, the combined population is sorted with their rank and crowding distance again to select elites as the new archive in the next generation. The details of AMGA2 can be found in [8].

Selecting the final solution from Pareto set. The R-MOSSC returns a set of solutions. To compare with other single objective optimization algorithms, a final solution is selected from the set according to their similarity of the partition matrix $\mathbf{U}$. The similarity is estimate by Normalized Mutual Information (NMI) [10]. Thus, the similarity measure is defined as follows:

$$sim(x) = \sum_{x' \in PS} NMI(\mathbf{U}_x, \mathbf{U}_{x'}) \tag{8}$$

The solution in PS with the highest score would be chosen as the final solution.

3 Experimental Study

To test the effectiveness of R-MOSSC, it is compared with PSOVW [6], EWKM [3], ESSC [4], KM [5] and R-KM [7]. We use NMI [10] as the performance metrics. All the experiments are run on PC (2.66 GHz CPU and 4 GB RAM) based on WEKA Data Mining Platform [11].

R-MOSSC and PSOVW are population based algorithms while the others are single point based methods. To evaluate the performance of these algorithms fairly, the maximum number of function evaluations (MaxFES) is used as terminal condition for all considered algorithms. The MaxFES is set to 500. For R-MOSSC, the initial population size and the archive size are set to 20. The population size of PSOVW is set to 20. Other parameter settings are shown in Table 1. The parameters without mentioning here are set to the same values used in the original literatures [4, 6, 8].

Eight data sets include UCI data sets (http://www.cs.waikato.ac.nz/∼ml/weka/), biomedical data sets (http://datam.i2r.a-star.edu.sg/datasets/krbd/) and the NIPS feature selection challenge data sets (http://www.nipsfsc.ecs.soton.ac.uk/datasets/) are used to evaluate the performance of R-MOSSC. The details of these data sets are shown in Table 2. We run R-MOSSC, KM and R-KM on each data set for 30 trials. Note that, PSOVW, EWKM and ESSC are parameter-sensitive. The suitable parameter settings for each data set are different. Thus, for each data set, we run these algorithms 30 times with every parameter setting shown in Table 1. The best average results are reported. All results are shown in Tables 3 and 4. The best result of each data set is shown in bold.

From Table 3, we can find: (1) the performance of R-MOSSC is better than that of other algorithms in most cases. It archives better average results on 7 data sets. Note that the results of PSOVW, EWKM and ESSC are obtained by carefully adjusting their parameter settings for different data sets. That is, it is difficult to find suitable settings for them without costly running and tuning; (2) the standard deviation of results of R-MOSSC is lower than those of other algorithms. In conclusion, R-MOSSC is better and more stable than other algorithms.

Table 4 shows the running time of all algorithms on each data set. R-MOSSC needs 1–3 times of the running time of PSOVW, and need 1–2 times of the running time of RKM. The reason is that R-MOSSC spends more computation resource for

Table 1 Parameter settings of algorithms

Algorithm	Parameter settings
PSOVW	$\beta = 2, 3, 4, 5, 6, 7, 8$
EWKM	$\gamma = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.1, 1.2, 1.3, 1.4,$ 1.5, 2, 3, 4, 5, 6, 7
ESSC	$\eta = 0, 0.01, 0.05, 0.1, 0.2, 0.3, 0.5, 0.7, 0.9$
	$\gamma = 1, 2, 5, 10, 50, 100, 1,000$
RKM	$\alpha = 3$

Table 2 Summary of data sets

Data set	Size of data set	Number of dimensions	Number of clusters
Balance scale	625	4	3
Ecoli	336	7	8
Heart statlog	270	13	2
Mfeat morphological	2000	6	10
Leukemia	72	7129	2
Breast cancer	97	24481	2
Leukemia stjude	327	12558	7
Arcene	200	10000	2

Table 3 Comparisons of different algorithms (NMI)

Algorithm		PSOVW	EWKM	ESSC	KM	RKM	R-MOSSC
Balance scale	Avg	0.1135	0.0982	0.1142	0.1159	0.1049	**0.1198**
	Std	0.0129	0.0185	0.0278	0.0560	0.0334	0.0142
Ecoli	Avg	0.4835	0.5915	0.5983	0.5926	0.5890	**0.6152**
	Std	0.0529	0.0277	0.0172	0.0281	0.0202	0.0119
Heart statlog	Avg	0.1695	0.1146	0.2214	0.2013	0.2394	**0.2644**
	Std	0.0680	0.0770	0.1049	0.1112	0.0601	0.0097
Mfeat morphological	Avg	0.6423	0.6493	0.6751	0.6513	0.6615	**0.6760**
	Std	0.0319	0.0331	0.0153	0.0353	0.0356	0.0050
Leukemia	Avg	**0.1410**	0.0653	0.1234	0.0998	0.0965	0.1382
	Std	0.0387	0.0493	0.0358	0.0523	0.0492	0.0346
Breast cancer	Avg	0.0393	0.0365	0.0465	0.0373	0.0477	**0.0740**
	Std	0.0028	0.0086	0.0379	0.0181	0.0103	0.0494
Leukemia stjude	Avg	0.4027	0.1674	0.4396	0.2843	0.2962	**0.4579**
	Std	0.0748	0.0896	0.0622	0.1152	0.1109	0.0833
Arcene	Avg	0.0819	0.0697	0.0766	0.0680	0.0680	**0.0823**
	Std	0.0012	0.0279	0.0290	0.0241	0.0287	0.0000

Table 4 Running time (second) of algorithms performed on data sets

Algorithm	PSOVW	EWKM	ESSC	KM	RKM	R-MOSSC
Balance scale	0.0517	0.0687	1.3963	0.0400	0.0540	0.0783
Ecoli	0.0580	0.0840	0.8930	0.0413	0.0600	0.1097
Heart statlog	0.0557	0.0733	1.4590	0.0280	0.0400	0.0633
Mfeat morphological	0.2773	0.4107	14.1370	0.2330	0.3800	0.6253
Leukemia	6.8583	9.4153	102.9800	2.6217	5.9040	7.5720
Breast cancer	33.3903	48.3860	317.9820	16.3483	31.8280	41.4897
Leukemia stjude	60.9857	104.7787	4365.9800	42.4197	88.2400	115.2667
Arcene	13.4593	40.3863	183.1827	13.8207	32.6520	37.5067

non-dominated sorting and crowding distance assignment to deal with multi-objective optimization. Benefiting from its filter framework and search strategy, R-MOSSC can make up the computation cost for non-dominated sorting and

crowding distance assignment with its simpler updating of **W** and **V**. Thus the running time of R-MOSSC is equal to that of EWKM. The ESSC is much slower than other algorithms because of the computation of fuzzy partition. KM is the simplest one and thus the fastest one. Note that, PSOVW, EWKM and ESSC need to tune parameter settings by running them multiple times with different parameter values for different data set. The running time shows in Table 4 just the average time of 30 runs producing the best average results (shown in Table 3) with the most suitable parameter setting. Consider the total time cost of multiple running with all parameter settings (i.e., the running time reported in Table 4 multiplied by the number of parameter settings shown in Table 1), the really running time of PSOVW, EWKM and ESSC is much longer than that of R-MOSSC.

4 Conclusion

This paper proposes a soft subspace clustering algorithm combining the advantages of MOEA and filter approach. Thus, the propose algorithm, called R-MOSSC, is a filter framework based on MOEA. Simulation results show that R-MOSSC is significantly better than the previously existing algorithms. This study will be further extended to real-world applications, such as text clustering and image segmentation. In addition, automatic clustering is another problem to be solved in future.

References

1. Kriegel HP, Kröger P, Zimek A (2009) Clustering high-dimensional data: a survey on subspace clustering, pattern-based clustering, and correlation clustering. TKDD 3:1–58
2. Beyer KS, Goldstein J, Ramakrishnan R, Shaft U (1999) When is "nearest neighbor" meaningful? In: Beeri C, Buneman P (eds) ICDT '99. Springer, London, pp 217–235
3. Jing L, Ng MK, Huang JZ (2007) An entropy weighting k-means algorithm for subspace clustering of high-dimensional sparse data. IEEE T Knowl Data Eng 19:1026–1041
4. Deng Z, Choi KS, Chung FL, Wang S (2010) Enhanced soft subspace clustering integrating within-cluster and between-cluster information. Pattern Recogn 43:767–781
5. MacQueen JB (1967) Some methods for classification and analysis of multivariate observations. In: Cam LML, Neyman J (eds) Proceedings of the fifth berkeley symposium on mathematical statistics and probability, vol 1. University of California Press, pp 281–297

6. Lu Y, Wang S, Li S, Zhou C (2011) Particle swarm optimizer for variable weighting in clustering high-dimensional data. Mach Learn 82:43–70
7. Boongoen T, Shang C, Iam-On N, Shen Q (2011) Extending data reliability measure to a filter approach for soft subspace clustering. IEEE Trans Syst Man Cybern B 41:1705–1714
8. Tiwari S, Fadel G, Deb K (2011) AMGA2: improving the performance of the archive-based micro-genetic algorithm for multi-objective optimization. Eng Optim 43:377–401
9. Zhou A, Qu BY, Li H, Zhao SZ, Suganthan PN, Zhang Q (2011) Multi-objective evolutionary algorithms: a survey of the state of the art. Swarm Evol Comput 1:32–49
10. Strehl A, Ghosh J (2003) Cluster ensembles—a knowledge reuse framework for combining multiple partitions. J Mach Learn Res 3:583–617
11. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009) The WEKA data mining software: an update. SIGKDD Explor 11:10–18

A New Type of Wavelet Threshold Denoising Algorithm

Li Kun Xing, Sen Qi and Wen Jing Wang

Abstract According to spectrum subtraction, this paper puts forward a new type of threshold value determination algorithm. Firstly, through the artificial extraction or by the zero point detection method, extract background noise from no sound segment. Secondly, do wavelet decomposition with background noise, and determine the threshold value on the basis of each layer's wavelet decomposition coefficient. Then, we can make a speech enhancement for the speech signal with noise. The simulation results show that this algorithm can effectively remove the noise component and keep the details of the useful signal characteristics very well. More over, the amount of calculation is far less than the traditional threshold algorithm's.

Keywords Wavelet analysis · Threshold de-noising · Speech enhancement

1 Introduction

The basic idea of using wavelet threshold de-noising algorithm to make a speech enhancement is: when we are making wavelet decomposition to the speech signal with noise, speech signal's wavelet coefficient is concentrated in the dense area, and the absolute value of wavelet coefficients is relatively large; On the contrary, noise signal wavelet coefficient distributes relatively dispersive, so the absolute value is small [1]. Based on the features of different layers' wavelet decomposition coefficient, we can do the multi-scale wavelet transform with noise signal, then cut or completely remove wavelet coefficients of noise component for each layer's coefficient, meanwhile retain speech signal's wavelet coefficient; After that, we can carry through wavelet reconstruction, so as to achieve the effect of speech enhancement.

L. K. Xing (✉) · S. Qi · W. J. Wang
College of Electrical and Information Engineering, Anhui University of Science
and Technology, Huainan 232001 Anhui, China
e-mail: xinglikun@126.com

There is a critical value between useful signal and noise signal in the high frequency wavelet coefficient. On the basis of experience or some testimony, we can find out the critical value and set it to threshold value. Then, the next work is to do threshold function processing [2]. The ideal threshold is equal to the critical value, so that the noise component can be entirely eliminated. At the same time, the useful signal can also be well reserved. Further more, we can get pure useful signal just through reconstructing wavelet coefficient. However, the noise component's critical value is quite close to the useful signals in actual applications. At this time, it directly affects the noise reduction effect. If the threshold value is too small, it's unable to effectively remove noise component, but if too large, it will cause the signal distortion [3]. This paper refers to spectrum subtraction. Using the improved threshold value determination algorithm, we can reduce relevant calculation by a large margin. In the meantime, it can effectively remove the noise component and keep the details of the useful signal characteristics quite well [4].

2 Spectrum Subtraction Speech Enhancement Technique

The spectral subtraction, which was put forward as a kind of speech enhancement algorithm, is partly based on the noise component statistics stability [5]; on the other hand, it's based on the irrelevance between additive noise and speech signal. The algorithm does not use the reference noise source, so we assumes that noise statistical is smooth, that is to say, the amplitude spectrum of noise component in speech and the noise amplitude spectrum of no sound segment in voice intermittent are expected to be the same [6]. We could take the noise spectrum estimation as the noise component spectrum. And noise speech spectrum minus noise spectrum equals the enhanced speech spectrum estimation.

The speech signal with noise is set to:

$$y(n) = s(n) + d(n), \quad 0 \le n \le N - 1 \tag{1}$$

where $s(n)$ is clean speech signal, and $d(n)$ is additive stationary noise. Thus, $y(n)$ is the speech signal with noise which has already done a adding-windows process. According to the short-term stability of speech signal, the noise signal needs the adding-window process when we use spectrum subtraction to do a speech enhancement. Set $d(n)$ to Gaussian white noise, and its Fourier transform still meet the Gaussian distribution. The basic idea of spectral subtraction is to take the spectrum of the speech signal with noise to subtract the noise spectrum extracted from no sound segment, and then do the inverse Fourier transform.

Doing Fourier transform with y(n), $s(n)$ and $d(n)$, we can get Y_k, S_k and N_k. Hence:

$$Y_k = S_k + N_k \tag{2}$$

$$|Y_k|^2 = |S_k|^2 + |N_k|^2 + S_k N_k^* + S_k^* N_k \tag{3}$$
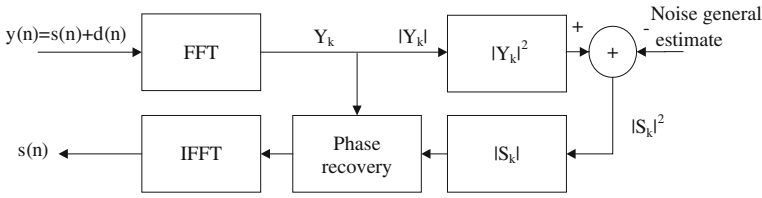
Fig. 1 The system structure figure of spectral subtraction

Because the noise component and the speech signal are not related, S_k and N_k distribute independently. N_k is a Gaussian distribution. Hence:

$$E||Y|^2|_k = E||S_k|^2| + E||N_k|^2| \tag{4}$$

According to the estimated noise spectrum in no sound segment, we can get the estimated value of the clean speech signal:

$$|Y_k|^2 = |S_k|^2 + \lambda_n(k) \tag{5}$$

$$|\hat{S}_k| = \left[|Y_k|^2 - E|N_k|^2\right]^{1/2} = \left[|Y_k|^2 - \lambda_n(k)\right]^{1/2} \tag{6}$$

where $|\hat{S}_k|$ is the enhanced speech signal spectrum amplitude.

The principle of the spectrum subtraction is shown in the Fig. 1:

The advantage of spectrum subtraction is that the algorithm is simple, and easy to realize. We can do it only through positive and negative Fourier transform. In the actual application, it doesn't need too many hardware resources [7]. Its disadvantage is unable to do speech enhancement effectively with low signal noise ratio (SNR) signal. If the estimation value of noise spectrum is too large, it will cause significant distortion to the speech signal. At the same time, it will join in a sense of rhythm, similar to the music's residual noise [8]. That is why this algorithm called spectrum subtraction.

3 Improved Threshold Value Selection Algorithms

This paper refers to spectrum subtraction, and then puts forward an improved threshold value determination algorithm. First of all, through the artificial extraction or by the zero point detection method, the system can extract background noise from no sound segment. Secondly, it will do wavelet decomposition with background noise, and determine the threshold value on the basis of different layers' wavelet decomposition coefficient. Finally, this algorithm could make a speech enhancement for the speech signal with noise. The expression of threshold algorithm is:

$$T_j = \max\left(|d_{j,k}|\right) \tag{7}$$

where j is the number of decomposition layer, and $\max\left(|d_{j,k}|\right)$ is the absolute value of the maximum of wavelet coefficient on different decomposition layers.

Now we do noise reduction processing with bumps and blocks separately, using this threshold value determination criterion. The signal-to-noise ratio (SNR) of bumps is 21.1208 dB, while the blocks' SNR is 19.7588 dB. The effect of noise reduction is shown in Fig. 2:

As shown in Fig. 2, the improved threshold algorithm can effectively remove the noise component, and keep the details of the useful signal characteristics quite well. Using the improved threshold algorithm, we can do experiments with a series of bumps in different SNR, and observe the corresponding noise reduction effects. After that, we should calculate the new SNR under the improved threshold algorithm and the heuristic threshold value criterion. The results are shown in Table 1:

Analyzing figures in Table 1, we know obviously: after noise reduction, the SNR of the speech signal with noise has a certain degree of ascending. The effects of ascending SNR are roughly equal under the two kinds of methods. And the SNR of improve threshold algorithm is slightly low. Along with the SNR decreasing, the gap to ascend SNR is increasing. However, the SNR can get ascended obviously. Improved threshold algorithm's advantage lies in the amount of calculation is far less than the heuristic threshold criterion's.

Analyzing figures in Table 1, we know obviously: after noise reduction, the SNR of the speech signal with noise has a certain degree of ascending. The effects of ascending SNR are roughly equal under the two kinds of methods. And the SNR of improve threshold algorithm is slightly low. Along with the SNR decreasing, the gap to ascend SNR is increasing. However, the SNR can get ascended

Fig. 2 Noise reduction effect using hard threshold function from the improved threshold

Table 1 Comparison of the noise reduction between heuristic and improved threshold (dB)

Before noise reduction	Improved threshold algorithm	Heuristic threshold algorithm
27.1414	31.9209	31.7317
21.1208	26.1535	26.5888
17.5990	23.5319	23.6970
11.5784	18.7629	18.3326
7.1414	14.2164	14.5177
−0.4628	8.1729	8.3257
−6.8380	2.3517	3.5164
−10.9204	−1.3616	−0.2077

obviously. Improved threshold algorithm's advantage lies in the amount of calculation is far less than the heuristic threshold criterion's.

Analyzing figures in Table 1, we know obviously: after noise reduction, the SNR of the speech signal with noise has a certain degree of ascending. The effects of ascending SNR are roughly equal under the two kinds of methods. And the SNR of improve threshold algorithm is slightly low. Along with the SNR decreasing, the gap to ascend SNR is increasing. However, the SNR can get ascended obviously. Improved threshold algorithm's advantage lies in the amount of calculation is far less than the heuristic threshold criterion's.

Four improved threshold value method in the application of speech enhancement.

Now we use the improved threshold value method to do threshold value noise reduction, and observe the corresponding effects of speech enhancement. Speech signal derives from the matching Yu Yin Ku of a book, common the computer-based signal processing practice, which is written by J.H.M McClellan, A.V.O Oppenheim and so on. A period of English speech "I like," is perfect for speech enhancement experiment, as the voice contains surd and dullness. And it has certain representativeness in enhancement experiments (Figs. 3, 4, 5).

From the above experiment, we know the improved threshold value determination algorithm can have a good speech enhancement effect, no matter using soft threshold function or hard threshold function. Using soft threshold value, the enhanced speech signal will be smoother and its mutation component will be less,

Fig. 3 Original speech signal
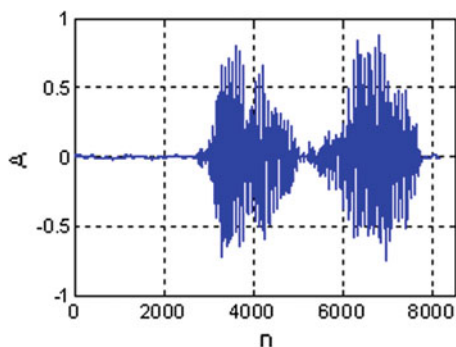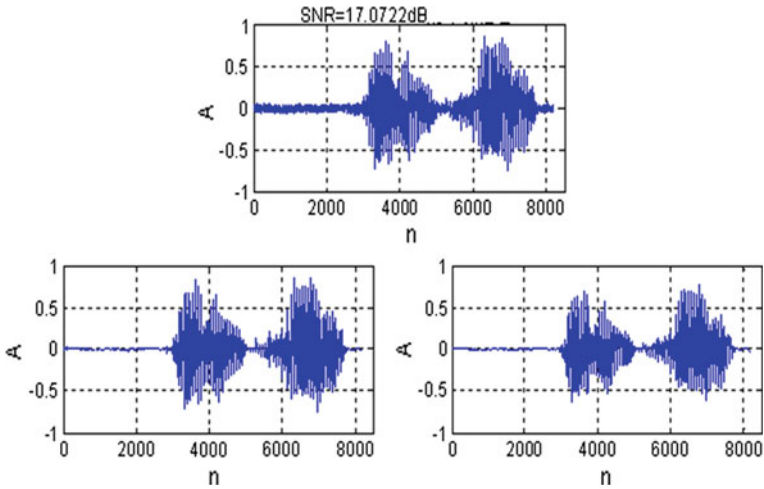
Fig. 4 Noise reduction effect using improved threshold *SNR = 17.0722 dB*
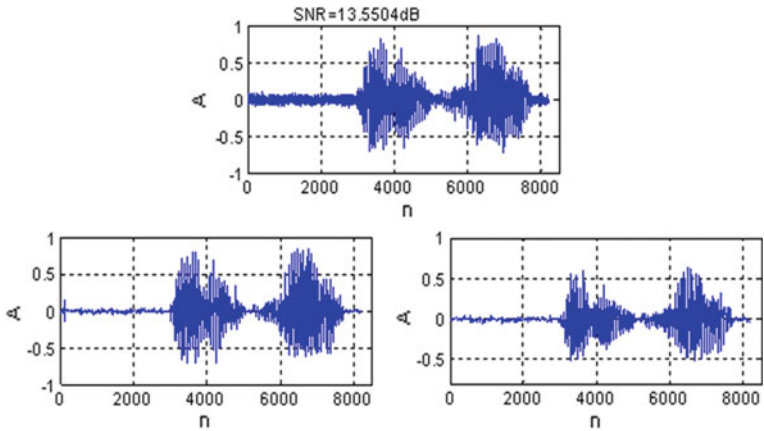
Fig. 5 Noise reduction effect using improved threshold *SNR = 13.5504 dB*

but it will change the signal amplitude. Human ear is more sensitive to frequency rather than signal amplitude, so the change of amplitude does not have a great influence to the sound effects of speech signal. When the signal-to-noise ratio is lower than zero, noise interference will be large and the enhanced speech signal will still remain some noise. However, the speech signal with noise compared to before has already got an obvious improvement.

4 Conclusion

This paper, basing on spectrum subtraction, puts forward an improved threshold value determination algorithm. Its advantage lies in that the amount of calculation is less than the traditional threshold algorithm's. Besides, this algorithm can effectively remove the noise component, as well as keep the details of the useful signal feature very well.

References

1. Chen S, Wang J (2004) Speech enhancement using perceptual wavelet packet decomposition and teager energy operator. J VLSI Sig Process 36:125–139
2. Daqrouq K, Abu-Isbeih IN,Daoud O,Khalaf E (2010) An investigation of speech enhancement using wavelet filtering method. Int J Speech Technol 13:101–115
3. Lu CT, Wang HC (2007) Speech enhancement using hybrid gain factor in critical-band-wavelet-packet transform. Digit Sig Process 17:172–188
4. Johnson MT, Yuan X, Ren Y (2007) Speech signal enhancement through adaptive wavelet thresholding. Speech Commun 49:123–133
5. Lu CT (2011) Enhancement of single channel speech using perceptual-decision-directed approach. Speech Commun 53:495–507
6. Yang GS, Liu YX, Jia Q (2011) Adaptive chaotic signal noise reduction based on double ascension wavelet. J Electron 1:13–17
7. Wang Z, Lv S, village Wu H (2008) The determination for the local singularity index and its spectrum of fractal function. J Electron Inf 30(2):290–292
8. Wang Z, Lv S, village Wu H (2008) A method to determine fractal function's singularity based on a wavelet spectrum. J Electron 4(3):482–485

The Application of DNA Nanoparticle Conjugates on the Graph's Connectivity Problem

Yafei Dong, Yanchai Wang, Jingjing Ma and Li Jia

Abstract A DNA computing algorithm is proposed in this paper which uses the assembly process of DNA-AuNP (DNA Au nanoparticle) conjugates to solve an NP-complete problem in the Graph theory, the connectivity problem, and a 3D DNA self-assembly algorithm model are also established. According to the algorithm we need to design the special DNA-AuNP conjugates which will assemble based on a specific graph, then a series of experiments are performed to get the final answer. This biochemical algorithm could reduce the complexity of the connectivity problem. The biochemical experimental technologies are mature and available, which will provide a practical way to validate the practicability and effect of DNA self-assembly algorithm model.

Keywords DNA computing · DNA-AuNP · Graph's connectivity · 3D model

Y. Dong (✉) · J. Ma (✉) · L. Jia
College of Life Science, Shaanxi Normal University,
Xi'an 710062 Shaanxi, People's Republic of China
e-mail: dongyf@snnu.edu.cn

Y. Dong · Y. Wang
College of Computer Science, Shannxi Normal University,
Xi'an 710062 Shaanxi, People's Republic of China

Y. Wang
e-mail: wangyanchai1025@163.com

J. Ma
School of Electronics Engineering and Computer Science, Peking University,
Beijing 100871, People's Republic of China
e-mail: casy-20@126.com

1 Introduction

The research of the nanotechnology has penetrated into many fields [1]. DNA is used to action as a variety of nanocomponents because of its unique physico-chemical property. DNA nanotechnology can [2, 3] not only provides the new methods and techniques for nanocomponents manufacturing and the research on the electronic components at molecular level, but also has its special value in DNA computer, nanobiological equipment, gene therapy and so on.

Based on DNA assembly which is a widespread formation system in nature, we construct nano components by using DNA [4–7]. This is known to all, cells are constituted by all sorts of biomolecule and organisms control physiological activities through biochemical manner. Molecular computing that serves as a bridge between biochemical processes and microprocessor-controlled mechanical components, and designs a new computing model called biochemistry algorithm.

DNA self-assembly [8] offers a route to fabrication with subnanometre precision of complex structures from simple components, for example, two complementary single-stranded DNA can cross and form a double-stranded DNA according to the Watson–Crick base pair without any external factors. This process is driven by a series of non-covalent interactions, such as electrostatic force, hydrophobic force, Van der Waals force, hydrogen bond and so on. Therefore, we can conduct molecular computing research by using DNA self-assembly mechanism under the appropriate laboratory condition.

Self-assembly has been getting increased attention as a main route of nanotechnology. Sharma et al. [9] used DNA-AuNP conjugates self-assembly to construct a 3D tubular structure of AuNP which can form different structures like single, double or even nested helixes by controlling AuNP size. Mastroianni and Claridge [10] reported the method to building 3D structure of AuNP. They established a pyramid structure through the AuNP self-assembly decorated by DNA strand. The AuNP size is adjusted to build a 3D chiral structure in pyramid structure's four vertices.

The goal of this paper is to present a DNA computing method using DNA-AuNPs conjugates self-assembly to solve an NP complete problem on connectivity graph theory. According to the algorithm design, the goal may be approached in two steps: first, we should design the conjugates of DNA-AuNPs self-assembly on the basis of specific connectivity problem; and second, solve the connectivity problems through a series of experiments. The proposed method reduce the computing complexity compared to traditional electronic computer.

2 DNA Nanoparticles Conjugates

The information capacity of biomolecules can be used to guide nanocrystal to form the 3D structure because of its unique identification characteristics. It suggests that nanocrystal and biomolecules conjugates have a good development prospects in

biological diagnosis and nanotechnology. There are many feasible ways to get the nanocrystal and biomolecules conjugates, such as biotin, avidin, antigen protein, antibody protein, polypeptide, protein etc.

A 12-base oligonucleotide is in stable thermodynamics state at room temperature, the length of which is approximately to 4 nm, or the size of a nanocrystal, and enough for completing pre-designed self-assembly by a DNA strands guided nanoparticle. This remarkable property has been used in several fields including creating 3D aggregates of nanocrystals [11], attaching nanocrystals to surfaces [12], creating small groupings of nanocrystals [13–15]. The common feature of DNA-AuNP conjugates used in series of studies is connection of tens of thousands of DNA strands on the AuNPs surface, and then implements the designed self-assembly, as shown in Fig. 1.

It is critical for diagnostics to control over the precise number of oligonucleotides per nanocrystal, because we often need quantify the number of hybridization events, rather than just assess the presence of a particular sequence. As shown in the Fig. 2, different numbers of single-stranded oligonucleotides from AuNPs have been separated by Daniela et al. [16], which provides the necessary conditions for research on nanocrystal and DNA conjugates. Nanocrystals bearing different numbers of oligonucleotides could serve as elementary construction units: vertex (4 strands), corners (3 strands), lines (2 strands), terminus (1 strand), etc. The nature of interaction between the polyanionic oligonucleotides is far from understood but is a very interesting problem in biophysics.

Daniela Zanchet demonstrated the electrophoretic isolation of discrete DNA-AuNP conjugates. We have chosen AuNPs because they are stable and processable in the high ionic strength buffers needed for manipulation of DNA. By adjusting the DNA: Au ratio, we can control the average number of DNA strands per nanoparticle, but there will always be a distribution of oligonucleotides present. Thus, we need a special technique for separation and isolation of the conjugates that is very sensitive to the number of bound strands.

Gel electrophoresis [17] is an important technique in biology and is widely applied in discrete groupings of AuNPs connected with DNA. In this technique, charged particles migrate in a gel under an electric field; particle mobility depends on their charge and size. Electrophoresis is also a useful tool in colloidal science, and it was recently applied to isolate and research on the characteristic of DNA-AuNPs conjugates. As shown in the Fig. 3, We can use DNA to control the number and size of nanoparticles linked, the distance between nanoparticles, the dimensionality and the flexibility of the structures. AuNPs have homogeneity of

Fig. 1 DNA-AuNP conjugates
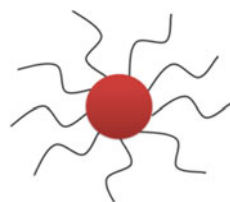
Fig. 2 Nanocrystal connecting different numbers of oligonucleotide

Fig. 3 Electrophoretic mobility of DNA-AuNPs in the gel electrophoresis
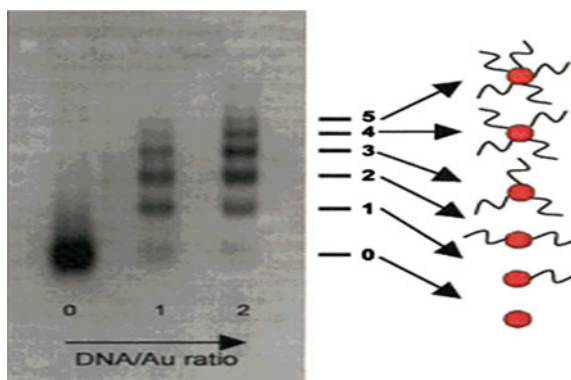
the size and shape, and they are also easy to manipulation in the solution. So AuNps can be extended to several other systems.

3 Connectivity of Graphs

Connectivity is an NP-complete problem in the graph theory. It is very important for the construction of reliable communications network, because it is closely related with the network model and combinational optimization. Therefore, it has important theory significance and the strong application value.

For example, let $G = (V, E)$ be a non-trivial connected graph. If V' is a subset of V and G-V' is disconnected. V' is vertex cutset of G. If vertex cutset contains k elements, we can use k-vertex cutset. Complete graph don't have vertex cutset. If G has at least different pairs of non-adjacent vertices, k is G's connectivity and is the minimum size of a vertex cutset in graph G. So the vertex-connectivity is the minimum size of a vertex cutset, we can use $\kappa(G)$ to denote connectivity of the graph G. If E' is subsets of E and G-E' is disconnected, E' is known as the edge cut of G. If edge cut contains k elements, we can use k-edge cut. So edge-connectivity is the minimum cardinality among all edge-cuts in the graph G. The edge-connectivity is deleted edge containing the least numbers, such that G is not connected. We can use $\lambda(G)$ to denote. For the graph G often has the following inequality formula [18, 19]

$$\kappa(G) \leq \lambda(G) \leq \delta(G) \tag{1}$$

Fig. 4 G
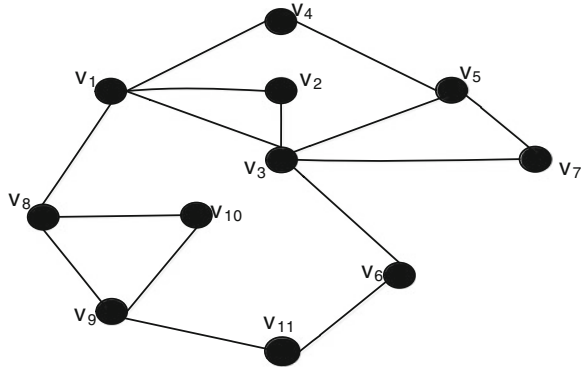
$\delta(G)$ is the minimum degree of the vertex in the graph G. Towards to proof of the theorem and the concept of graph theory is undefined in the paper, readers can look for reference [20]. Solving the connectivity issues is currently not the efficient algorithm, which mainly uses maximum flow algorithm to solve edge-connectivity. However, there is not a good algorithm to solve the vertex connectivity [21]. Gang et al. [22] designed a biological algorithm with a 3D structure of DNA to solve the connectivity, Shengmin [23] designed a biological evolutionary algorithm with a similar structure.

As indicated in Fig. 4, if G = (V, E) is an undirected graph, V is the set of vertices, E is the set of edges. G is stored in a matrix, as indicated in Fig. 5. 1 means that both nodes have access to, 0 means that both nodes don't have access to in Fig. 5. And each row (column) of the matrix represent the degree of the nodes, so we obtain that the degree of V_2, V_4, V_6, V_7, V_{10} and V_{11} are 2 in the graph G, degree of V_5, V_8 and V_9 are 3, degree of V_1 is 4, degree of V_3 is 5. Connecting two single strands structure of DNA conjugates represents V_2, V_4, V_6, V_7, V_{10} and V_{11}

Fig. 5 Matrix storage of the graph G

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$
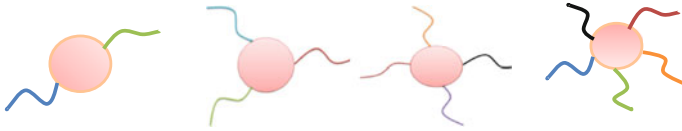
Fig. 6 Four kinds of vertex structure

in the Fig. 6, connecting three single strands structure of the DNA conjugates represents V_5, V_8 and V_9 in the Fig. 6, connecting four single strands structure of the DNA conjugates represents V_1 in the Fig. 6. Connecting five single strands structure of the DNA conjugates represents V_3 in the Fig. 6.

4 Computing Model of DNA Nanoparticles Conjugates Self-assembly

Using self-assembly computing model of DNA nanoparticles conjugates to solve connectivity, we need to use DNA-AuNPs conjugates. According to the vertices and edges of the specific graph, we design the specific number and sequence of bases in DNA-AuNPs conjugates, which is used to form the desired structure for completing self-assembled.

In the Fig. 4, for example. V_2, V_3 and V_4 are adjacent to V_1, V_1 and V_3 are adjacent to V_2. When we design DNA-AuNP conjugates which represent V_1 and V_2, the DNA-AuNP strands of V_1 and V_2 contain the complementary base sequences, and the other three strands of V_1 contain the complementary base sequences are to V_3, V_4 and V_8. By that analogy, the DNA-AuNPs conjugates not only can represent the information about vertex, but can precisely represent the information about edges. When two vertices are linked, the complementary sequences of DNA-AuNPs will take place hybridization reaction to form stable 3D nanostructure. This reaction is easy to be controlled in the appropriate temperature, PH and density of buffer solution.

According to the concept of connectivity, we use $\kappa(G)$ to denoted. If we delete V_1 and V_6, G is not connected. So connectivity of graph G is 2. The biological algorithm to solve connectivity is completed by a series of biochemical experiments in this paper.

First, pre-designed DNA-AuNPs can produce 3D nanostructure of G through annealing reaction.

Second, we need to delete the vertex. We delete vertex through strand displacement reaction. For example, if we want to delete the V_6, we must design the DNA strand to displace two single-strands of the V_6. When this strand was added to 3D nanostructure of G under laboratory conditions, DNA-AuNP conjugates standing for V_6 will be replaced. V_6 will be deleted from the reaction. And so on, 11 vertices of graph G need to design the substituted strands. The substituting

strands should contain the more long sequence than the substituted strands in the conditions of the strand displacement reaction. We must consider in the design of substituting strands. Let V_3 and V_6 be an example, as indicated in Fig. 7 [24].

To get its connectivity in graph G, we delete vertices

$$2^{11} - 2 = 2{,}046 \qquad (2)$$

The space of the solution is relatively large, we can use the inequality formula (1) to narrow the solution space. According to the inequality formula, we can obtain $\delta(G) = 2$ from Fig. 5. That is to say, solution space will be reduced to

$$C_{11}^1 + C_{11}^2 = 66 \qquad (3)$$

So the solution space can be reduced by this theorem. Nanostructures contain correct solution in the 66 combinations after deleting vertices.

Third, detect the correct solution. We use the agarose gel electrophoresis. The different DNA-AuNPs conjugates present the different morphology in agarose gel electrophoresis, but they still follow the electrophoresis rules. For example particle mobility depends on their charge and size. As noted in the this paper and references [16], the conjugates of the DNA-AuNPs can distinguish in the electrophoresis experiments, the mobility become smaller when the numbers are more; the mobility of self-assembled containing maximum AuNPs are smallest in graph G. G_i was represented after i-vertices are deleted. We can use 'detect (G_i)' to denote. If G_i is not connected, the connected graph can be distinguished in the distribution of the electrophoresis and we let $t = 1$; if not, $t = 0$. Deleting V_1 and V_6, two kinds of the graph containing four and five AuNPs self-assembled are not connected. We can easily identify in the electrophoresis.

For other connectivity problem, biochemistry algorithm is feasible in this paper. Universal calculation of this algorithm is expressed as the following biological algorithm:

```
1 for i ← 1 to δ(G)
2 delete (G, i)
3 t = detect(Gᵢ)
4 if (t)
```


Fig. 7 Opening reaction of the strands connecting V_3 and V_6

TCAGCGATAC
CTATGCGACT
取代链：
TCATGATATGCTGA

TCAGCGATAC
CTATGCGACT
TCATGATATGCTGA
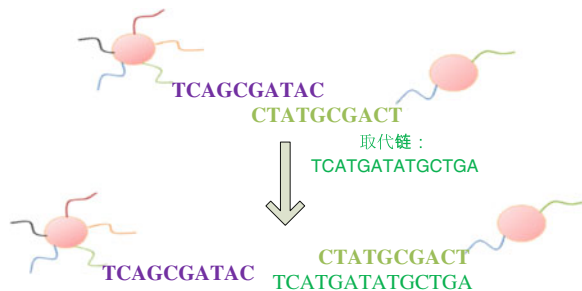
```
5     return κ(G) = i
6 else goto 1
7 end
```

5 Conclusion

In this paper, we come out with a way of using DNA-AuNPs conjugates self-assembly to solve connectivity problems and construct the 3D DNA self-assembly calculation model to solve the connectivity problem. Most important, the chemical method at design and synthesis of DNA-AuNPs conjugates is simple. The computing model mainly uses the programmability of base sequence in DNA and particular chemical structure and properties of nanoparticles. Compared to the tradition compute method at DNA, it not only reduces the complexities, but overcomes index exploration problem, some other detecting techniques can be well used to test the computational model in this paper such as transmission electron microscopy and fluorescence techniques.

Although this model is theoretical in this paper, we have reliable experiment and effective experiment technology. Nevertheless, we don't avoid the mistakes in the actual operation and experiments, so many factors are hard to control in the experiment. Potential steps in this direction, first, improving and testing theoretical model in the experiments. The next, making the theoretical model and experimental method better use in new technologies. Of course, these ways are efficient to establish DNA computing model with certain universality.

References

1. Chen CQ, Cui JZ (2011) Perspectives in mechanics of heterogeneous solids. Acta Mech Solida Sin 24(1)
2. Jing XH, Liu WQ, Chen JJ, Lin XQ (2007) Application of DNA nanotechnology. Prog Chem 19(4)
3. Xiao SJ (2011) The emerging field of RNA nanotechnology. Chem Life 2
4. Jing Y, Cheng Z (2008) Progress and difficulty in DNA self-assembly technology. Chin J Comput 31(12)
5. Song SB, Yin ZX, Zhen C, Hua C (2011) DNA self-assembly model for general satisfiability problem. J Chin Comput Syst 32(9)
6. Zhou YT, Li KL (2012) An algorithm for solving maximum clique problem based on self-assembly model of DNA. J Hunan Univ 39(9)
7. Zhang C, Yang J, Xu J (2011) Molecular logic computing model based on self-assembly of DNA nanoparticles. Chin Sci Bull 56(33)

8. Cheng Z, Jing Y, Jin X (2011) Logic calculation model on DNA self-assembly and nanoparticle molecular. Chin Sci Bull 56:2276–2282
9. Sharma J, Chhabra R, Cheng A, Brownell J, Liu Y, Yan H (2009) Control of self-assembly of DNA tubules through integration of gold nanoparticles. Science 313:112–116
10. Mastroianni AJ, Claridge SA (2009) Pyramidal and chiral groupings of gold nanocrystals assembled using DNA scaffolds. Jam Chem Soc 131:8455–8459
11. Mirkin CA, Letsinger RL (1996) A DNA-based method for rationally assembling nanoparticles into macroscopic materials. Nature 382:607–609
12. Taton TA, Mirkin CA, Letsinger RL (2000) Scanometric DNA array detection with nanoparticle probes. Science 289:1757–1760
13. Alivisatos AP, Johnsson KP, Peng XG, Wilson TE, Loweth CJ et al (1996) Organization of 'nanocrystal molecules' using DNA. Nature 382:609–611
14. Mucic RC, Storhoff JJ, Mirkin CA, Letsinger RL (1998) DNA-directed synthesis of binary nanoparticle network materials. J Am Chem Soc 120:12674–12675
15. Loweth CJ, Caldwell WB, Peng X, Alivisatos AP, Schultz PG (1999) DNA-based assembly of gold nanocrystals. Angew Chem Int Ed 38:808–1812
16. Daniela Z, Christine MM, Wolfgang JP et al (2001) Electrophoretic isolation of discrete Au nanocrystal/DNA conjugates. Nanoletters 1(1):32–35
17. Feng HY, Jing ZZ, Fang YX, Mo SK (2009) Two-dimensional electrophoresis and its application. Biotechnol Bull 1
18. Chen L (2000) The connectivity of vertex transitive graphs. J Xinjiang Univ 17(4)
19. Yang C, Xu JM (2008) Connectivity and edge-connectivity of strong product graphs. J Univ Sci Technol China 5
20. Bondy JA et al (1976) Graph theory with applications. London, pp 108–117
21. Shuhe W (2004) Graph theory. Science Press, Beijing, pp 141–144
22. Gang F, Shengmin Z, Jin X (2006) Three dimensional DNA graph structure solution to edge-connectivity. Syst Eng Electron 28(1)
23. Shengmin Z (2007) Three dimensional DNA structure solution to connectivity based on evolutionary algorithm. Comput Eng Appl 43(7)
24. Tingjie S, Haojun L (2012) Synchronized assembly of AuNPs driven by a dynamic DNA-fueled molecular machine. J Am Chem Soc 134:10803–10806

RETRACTED CHAPTER: Using Inner-Outer Factorization to Solve the Spectral Factor of Discrete-Time Descriptor Systems

Luhua Liang, Wei Xing and Rongwu Xiang

Abstract In this paper, we use the state-space realization of discrete-time descriptor system to solve the inner-outer and spectral factorization problems. The algorithm is based on finding two orthogonal matrices to decompose the pole separated realization of transfer function matrix, to get a stabilizing solution by sloving a algebraic Riccati equation which order usually smaller than the McMillan degree of the transfer function. We give a theorem to discuss the relation of inner-outer and spectral factorization and get the inner-outer factor of the system. Thus, the inner-outer factor is the spectral factor of the system. Finally, a simple numerical example is also illustrated.

L. Liang · R. Xiang (✉)
Staff Room of Mathematics,
Shenyang Pharmaceutical University, Shenyang 110016, China
e-mail: xrwlove@163.com

L. Liang
e-mail: liangluhua1988@163.com

W. Xing
Institute of Systems Science, Northeastern University,
Shenyang 11004, China
e-mail: awxing@21cn.com

1 Introduction

The problems of inner-outer factorization and spectral factorization are very important in the theory of control systems, especially in signal processing of robust filtering, linear quadratic optimal control, H_2 and H_∞ optimal control. In fact, inner-outer factorization is closely related to spectral factorization, we can calculate the outer factor of inner-outer factorization though the spectral factorization of the systems. The two methods have been used in many different files, such as in Ref. [1] give a iterative method about how to calculate the inner-outer factorization and spectral factorization, and which also give the convergence expression of iterative method. Ball had researched the inner-outer factorization in the continuous-time asymptotic stability of nonlinear state-space systems [2]. Anderson had given a spectral factorization of a class of matrices arising in tackling Wiener filtering theory and network synthesis via an algebraic procedure [3].

Studying the inner-outer factorization and spectral factorization often use two kinds of different ideas, one is based on a special or general method to calculate the problem, the other is moving the unstable zeros to the stable area by left multiplied with an all-pass factor in the systems. In this paper our method is based on finding a stable solution of algebra Riccati equation which rank is less than the McMillan degree of descriptor systems. And give a general solution for the arbitrary rank, zeros on imaginary axis or at infinity of transfer function.

2 The Description of Problem

Considering the discrete-time descriptor system [3]:

$$\begin{cases} Ex(k+1) & = Ax(k) + Bu(k) \\ y(k+1) & = Cx(k) + Du(k) \end{cases} \tag{1}$$

where, matrix E is singular, as $\det E = 0$. A, B, C, D are real matrix respectively, which are $n \times n, n \times m, p \times n, p \times m$. $A - zE$ is regular, i.e., it is square and $\det(zE - A) \neq 0$. D is full column rank, the transfer function matrix of this discrete-time descriptor system with minimal realization are defined as

$$G(z) = \left[\begin{array}{c|c} A - zE & B \\ \hline C & D \end{array} \right] \tag{2}$$

Definition 1 [4] Given a square real rational matrix fuction $\Phi(z)$, where $\Phi(z) = \Phi^T(z^{-1})$, with bounded $\Phi(e^{j\omega}) \geq 0$ and nonnegative definite for $\omega \in [-\pi, \pi]$, then there exists a real rational matrix $\phi(z)$, such that

$$\Phi(z) = \phi^T(z^{-1})\phi(z) \tag{3}$$

And we call equality (3) is the spectral factorization of $\Phi(z)$, and call $\phi(z)$ is the spectral factor of $\Phi(z)$. Usually, we request [5] the $\phi(z)$ on the right half-plane (Res > 0) which is analyzing, and also request $\phi(z)$ has the constant rank or equally has the right inverse matrix $\phi_r^{-1}(z)$ which is analyzing on the right half-plane. For a rational matrix with real coefficients $\Phi(z)$, we define its conjugate $\Phi^*(z) = \Phi^T(z^{-1})$. If $\Phi^*(z) = \Phi(z)$, then the equality (3) can be considered as

$$\Phi(z) = \phi^*(z)\phi(z) \tag{4}$$

For this spectral factorization problem, Crisrtian had given two different solution (see ref. [6]), Ridden and Anderson obtained a new result by doing a little amending based on the this. Anderson [7] gets a solution about (4) is an algebra method, which by solve a quadratic matrix equation (only include the constant matrix), the solution of (4) is the constant rank on the right half-plane. It can also be decided by solving satisfied other inequality of eigenvalue of matrix. If $\Phi(\infty)$ is a singular matrix, then must be solve some correlative calculation problem. The main characteristic of Anderson's algorithm is reported the constant matrix of quadratic equation to determine the eigenvalue is very useful.

In this paper, we want to use another method to get the spectral factor of system, which is very different from Anderson. We are going to use the inner-outer factorization to obtain the spectral factor. Following, there is the definition of inner-outer factorization.

Definition 2 [7] Inner-outer factorization is a real rational matrix $G(z)$ with stable decomposition and row full rank. It can be a product of a suitable inner factor and a stable MP (minimum phase) outer factor, such as

$$G(z) = \theta(z)\phi(z) \tag{5}$$

If $\theta(z)$ is stable and $\theta^*(z)\theta(z) = I$, then called $\theta(z)$ as inner factor. If $\phi(z)$ is stable and minimum phase, then called $\phi(z)$ as out factor.

More precisely, we consider the following spectral factorization problem: Given an arbitrary real rational $p \times m$ matrix $G(z)$, finding a rational matrix $\phi(z)$, such that

$$G^*(z)G(z) = \phi^*(z)\phi(z) \tag{6}$$

where $\phi(z)$ is the spectral factor, and (9) is the spectral factorization of $G^*(z)G(z)$.

3 Main Conclusion

Since our goal is the spectral factorization of $G(z)$ via state-space methods we are interested, therefore in spectral decompositions of the pole and system pencils with respect to the partition $\overline{C} = \overline{D} \cap D^c$, where D is an open unit disk, $\bar{D}$ is accordingly closed unit disk, C is the complex plane $\bar{C} = C \cup \{\infty\}, D^c = \bar{C} - \bar{D}$. Without restricting generality we may assume directly that $G(z)$ is given by a pole separated realization of the form;

$$G(z) = \left[\begin{array}{cc|c} A_a - zE_a & A_{as} - zE_{as} & B_a \\ 0 & A_s - zE_s & B_s \\ \hline C_a & C_s & D \end{array}\right] \tag{7}$$

where $\wedge(A_S - zE_s) \subset \overline{D}, \wedge(A_a - zE_a) \subset D^c$. A stable realition (2) always can separated of the form (7) by using spectral factorization on poles metric pencil $A - zE$. If the realization to start with has real coefficients, one can always determine a separated realization with real coefficients as well. In order to get the conclusion of this paper, firstly, give the next theorem.

Theorem 1 [8, 9] *Let $G(z)$ be a real rational matrix given by a pole separated realization* (7). *Then there exist two unitary matrices U and Z such that*

$$\begin{bmatrix} U & 0 \\ 0 & I \end{bmatrix} \left[\begin{array}{c|c} A - zE & B \\ \hline C & D \end{array}\right] Z = \left[\begin{array}{ccc|c} A_{r,-} - zE_{r,-} & * & * & * \\ 0 & A_{+,l} - zE_{+,l} & B_{+,l} & * \\ 0 & 0 & 0 & B_n \\ \hline 0 & C_{r,l} & D_{+,l} & B_{n\times} \end{array}\right] \tag{8}$$

in which:

(1) *the pencil $A_{r,-} - zE_{r,-}$ has full row rank in D^c and $E_{r,-}$ has full row rank.*

(2) *$E_{+,l}$ and B_n are invertible the pencil $\begin{bmatrix} A_{+,l} - zE_{+,l} & B_{+,l} \\ C_{+,l} & D_{+,l} \end{bmatrix}$ has full column rank in $\overline{D}$, and the pair $[A_{+,l} - zE_{+,l} \quad B_{+,l}]$ is stabilizable.*

Once the two spectral decompositions of the pole and system pencil obtained, we have the coefficients of a Riccati equations, whose guaranteed solutions can be used to write down directly state space formulas for the spectral factor $\phi(z)$. The following theorem contains the main result of the paper.

Theorem 2 *Let $G(z)$ be a real rational matrix with a pole separated realization* (7), *and let U and Z be orthogonal matrices for which* (8) *holds, then we have:*

(1) *The discrete-time algebraic Riccati equation*

$$\begin{aligned} A_{+,l}^T X A_{+,l} - E_{+,l}^T X E_{+,l} - (A_{+,l}^T X B_{+,l} + C_{+,l}^T D_{+,l}) (D_{+,l}^T D_{+,l} \\ + B_{+,l}^T X B_{+,l})^+ (B_{+,l}^T X A_{+,l} + D_{+,l}^T C_{+,l}) + C_{+,l}^T C_{+,l} = 0 \end{aligned} \tag{9}$$

has a stabilizing solution X_s such that $R = D_{+,l}^T D_{+,l} + B_{+,l}^T X B_{+,l} > 0$, and $\wedge(A_{+,l} + B_{+,l} F_s - zE_{+,l}) \subset D$ where $F_s = -(D_{+,l}^T D_{+,l} + B_{+,l}^T X B_{+,l})^+ (B_{+,l}^T X A_{+,l} + D_{+,l}^T C_{+,l})$ is the stabilizing Riccati feedback.

(2) Let X_s and F is the stabilizing solution and feedback matrix of Riccati equation (9), and $R = D_{+,l}^T D_{+,l} + B_{+,l}^T X_s B_{+,l}$, define

$$\phi(z) = \left[\begin{array}{c|c} A - zE & B \\ \hline \tilde{C} & \tilde{D} \end{array}\right] \tag{10}$$

$$\theta(z) = \left[\begin{array}{c|c} A_{+,l} + B_{+,l}F - zE_{+,l} & B_{+,l}R^{-\frac{1}{2}} \\ \hline C_{+,l} + D_{+,l}F & D_{+,l}R^{-\frac{1}{2}} \end{array}\right], \begin{bmatrix} \tilde{C} & \tilde{D} \end{bmatrix} = R^{\frac{1}{2}}\begin{bmatrix} 0 & -F & I & 0 \end{bmatrix}Z^T$$

$$\tag{11}$$

Then $\phi(z)$ is a spectral factor of $G(z)$, i.e., $\phi(z)$ is stable, minimum phase spectral factor and satisfies $G^*(z)G(z) = \phi^*(z)\phi(z)$.

Proof

(1) From the conclusion 2 in theorem, we have the matrix $E_{+,l}$ is invertible. Then we use $(E_{+,L}^{-1})^T, E_{+,l}^T$ difference premultiplied on the right and left of Riccati Eq. (9). we have

$$(E_{+,l}^{-1})^T A_{+,l}^T X A_{+,l} E_{+,l}^{-1} - X - [(E_{+,l}^{-1})^T A_{+,l}^T X B_{+,l} + (E_{+,l}^{-1})^T C_{+,L}^T D_{+,l}](D_{+,l}^T D_{+,l}$$
$$+ B_{+,l}^T X B_{+,l})^+ (B_{+,l}^T X A_{+,l} E_{+,l}^{-1} + D_{+,l}^T C_{+,l} E_{+,l}^{-1}) + (E_{+,l}^{-1})^T C_{+,l}^T C_{+,l} E_{+,l}^{-1} = 0$$

have change it, we can get

$$X = (E_{+,l}^{-1})^T A_{+,l}^T X A_{+,l} E_{+,l}^{-1} - [(E_{+,l}^{-1})^T A_{+,l}^T X B_{+,l} + (E_{+,l}^{-1})^T C_{+,l}^T D_{+,l}](D_{+,l}^T D_{+,l}$$
$$+ B_{+,l}^T B_{+,l})^+ (B_{+,l}^T X A_{+,l} E_{+,l}^{-1} + D_{+,l}^T C_{+,l} E_{+,l}^{-1}) + (E_{+,l}^{-1})^T C_{+,l}^T C_{+,l} E_{+,l}^{-1}$$

where $A_{+,l}^T = (E_{+,l}^{-1})^T A_{+,l}^T, \tilde{C}_{+,l} = C_{+,l}E_{+,l}^{-1}$. Form some character of Riccati equation, we know the above equation has unique solution and has a stable solution X_s, such that

$$R = D_{+,l}^T D_{+,l} + B_{+,l}^T X B_{+,l} > 0 \text{ and } R = D_{+,l}^T D_{+,l} + B_{+,l}^T X B_{+,l} > 0,$$

where $R = D_{+,l}^T D_{+,l} + B_{+,l}^T X B_{+,l} > 0$ is the stable Riccati feedback of system.

(2) Let

$$\begin{bmatrix} \hat{C} & \hat{D} \end{bmatrix} = R^{-\frac{1}{2}}\begin{bmatrix} \tilde{C} & \tilde{D} \end{bmatrix} = \begin{bmatrix} 0 & -F & I & 0 \end{bmatrix}Z^T \tag{12}$$

Using (10) and (11), we have

$$G(z) = \left[\begin{array}{cc|c} A_{+,l}+B_{+,l}F-zE_{+,l} & B_{+,l}\hat{C} & B_{+,l}\hat{D} \\ 0 & A-zE & B \\ \hline C_{+,l}+D_{+,l}F & D_{+,l}\hat{C} & D_{+,l}\hat{D} \end{array}\right] = \left[\begin{array}{cc|c} A_{+,l}+B_{+,l}F-zE_{+,l} & 0 & B_{+,l}D-X_1B \\ 0 & A-zE & B \\ \hline C_{+,l}+D_{+,l}F & C_e & D_{+,l}D \end{array}\right] = \left[\begin{array}{c|c} A-zE & B \\ \hline C_e & D_{+,l}\hat{D} \end{array}\right]$$

where $C_e = D_{+,l}\hat{C} + (C_{+,l} + D_{+,l}F)X_2$, matrix X_1 and X_2 satisfied the following equlity

$$\begin{bmatrix} X_1 & -B_{+,l} \end{bmatrix}\begin{bmatrix} A-zE & B \\ \hat{C} & \hat{D} \end{bmatrix} = (A_{+,l}+B_{+,l}F-zE_{+,l})\begin{bmatrix} X_2 & 0 \end{bmatrix} \tag{13}$$

because the matrix X_1 and X_2 are exist, we can have the following equality directly,

$$\begin{bmatrix} 0 & I & X_3 & -B_{+,l} \end{bmatrix}\begin{bmatrix} A_{r,-}-zE_{r,-} & * & * & * \\ 0 & A_{+,l}-zE_{+,l} & B_{+,l} & B_\times-zE_\times \\ 0 & 0 & 0 & B_n \\ \hline 0 & -F & I & 0 \end{bmatrix} = (A_{+,l}+B_{+,l}F-zE_{+,l})\begin{bmatrix} 0 & I & 0 & E_{+,l}^{-1}E_\times \end{bmatrix}$$

then $X_3 = [-B_\times + (A_{+,l}+B_{+,l}F)E_{+,l}^{-1}E_\times]B_n^{-1}$. We multiply the matrix Z^T on the above equation and use (8), (11), (12) we can have conclusion from (13) that

$$X_1 = \begin{bmatrix} 0 & I & X_3 \end{bmatrix}U\begin{bmatrix} X_2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & I & 0 & E_{+,l}^{-1}E_\times \end{bmatrix}Z^T \tag{14}$$

reversible change of the original system, we have

$$\begin{bmatrix} I & 0 \\ \bar{U} & 0 \end{bmatrix}\begin{bmatrix} A-zE & B \\ C & D \end{bmatrix} = \begin{bmatrix} A-zE & B \\ D_{+,l}\hat{C}+(C_{+,l}+D_{+,l}F)X_2 & D_{+,l}\hat{D} \end{bmatrix}$$

using matrix $\begin{bmatrix} U & 0 \\ 0 & I \end{bmatrix}$ left multiply on the above equality, right multiply with matrix Z, and use (8), (13), (14) and $\bar{U}U^T = \begin{bmatrix} U_{11} & U_{12} & U_{13} \end{bmatrix}$, we have

$$\begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \\ U_{11} & U_{12} & U_{13} & I \end{bmatrix} \begin{bmatrix} A_{r,-} - zE_{r,-} & * & * & * \\ 0 & A_{+,l} - zE_{+,l} & B_{+,l} & B_{\times} - zE_{\times} \\ 0 & 0 & 0 & B_n \\ \hline 0 & C_{+,l} & D_{+,l} & B_{n\times} \end{bmatrix} = \begin{bmatrix} A_{r,-} - zE_{r,-} & * & * & * \\ 0 & A_{+,l} - zE_{+,l} & B_{+,l} & B_{\times} - zE_{\times} \\ 0 & 0 & 0 & B_n \\ \hline 0 & C_{+,l} & D_{+,l} & (C_{+,l} + D_{+,l}F)E_{+,l}^{-1}B_{n\times} \end{bmatrix}$$

The above equality satisfied $U_{11} = 0, U_{12} = 0$ and $U_{13} = [-B_n + (C_{+,l} + D_{+,l}F)E_{+,l}^{-1}E_{\times}]B_n^{-1}$, then prove the existed of a matrix $\bar{U}$, and also give the proof of $G(z) = \theta(z)\phi(z)$.

4 Example

Consider the polynomial matrix

$$G(z) = \begin{bmatrix} z^2 + z + z & 4z^2 + 3z + 2 & 2z^2 - 2 \\ z & 4z - 4 & 2z - 2 \\ z^2 & z^2 - z & 2z^2 - 2z \end{bmatrix}$$

The structural elements of $G(z)$ are: a zero at $z = 1$; a pole at $z = \infty$; normal rank $r = 2$; unknown the spectral factorizations of $G(z)$. Using the relationship of inner-outer factorization and spectral factorization to work out the stable, minimum phase of $G(z)$.

Answer: An irreducible descriptor realization of $G(z)$ is given by

$$A - zE = \begin{bmatrix} -z & 0 & 1 & 0 \\ 0 & -z & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 4 & 2 \\ 0 & -1 & -2 \end{bmatrix},$$

$$C = \begin{bmatrix} 0 & 0 & -1 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, D = \begin{bmatrix} 1 & 2 & -2 \\ 0 & -1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$$

The minimum phase spectral factorization: With $Q = I$,

$$Z = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 2 & -1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 6 & 3 & 4 & -4 & -4 & -3 & 0 \\ -2 & -1 & -1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

we have

$$Q^{-1}\left[\begin{array}{c|c} A - zE & B \\ \hline C & D \end{array}\right]Z = \left[\begin{array}{c|c|c|c} A_{r,-} - zE_{r,-} & * & * & * \\ \hline 0 & A_{+,l} - zE_{+,l} & B_{+,l} & B_{\times} - zE_{\times} \\ \hline 0 & 0 & 0 & B_{n} \\ \hline 0 & C_{+,l} & D_{+,l} & B_{n\times} \end{array}\right] = \left[\begin{array}{cc|c|cc|c} 0 & 1 & -z & 2 & -1 & -1 & 0 \\ 0 & 0 & -z & -1 & 0 & -z \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{array}\right]$$

We solve the discrete-time algebraic Riccati equation (12) and get $X = 1$, $F^T = \begin{bmatrix} 0 & 0 \end{bmatrix}$, then the inner factor of inner-outer factorization is $\theta(z) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}z} & \frac{1}{\sqrt{2}} \end{bmatrix}^T$ and the minimum phase spectral factor is

$$\phi(z) = \begin{bmatrix} z^2 - z - 1 & -4z^2 - 3z - 2 & -2z^2 + 2 \\ \sqrt{2}z^2 & -\sqrt{2}(-4z^2 + z) & 2\sqrt{2}(z^2 - z) \end{bmatrix}.$$

5 Conclusions

We have presented a constructive state-space solution to the problem of inner-outer factorization and spectral factorization of the discrete-time descriptor system. An important feature of our approach is that we use orthogonal transformations to the largest extent possible. Firstly, we can always compute an irreducible pole separated realization of $G(z)$ by employing orthogonal transformations only. Further, the coefficients of the Riccati equation and the matrices U and Z could be obtained

by any staircase algorithm. The final step is to compute the solutions to standard Riccati equations by one of the existing numerically example solvers. This method gives a simple theory for calculating the spectral factor of a descriptor system. The conclusion can be extended to the J-spectral factorization problem.

References

1. Brunelli R, Poggio T (1993) Face recognition: feature versus templates. IEEE Trans PAMI 15(10):1042–1052
2. Ball JA, Peterson MA, Van der Schaft A (2004) Inner-outer factor relation for nonlinear noninvertible systems. IEEE Trans Autom Control 49(4):483–492
3. Anderson BDO (1967) An algebraic solution to the spectral factorization problem. IEEE Trans Automa Control AC-12: DD.410–414
4. Wang L, Zhang W, Liu Y (2007) The analytic computation of inner, outer and spectral factorization problems. J Shanghai Jiaotong Univ 41(6): 889–894 (in Chinese)
5. Oara C, Varga A (2000) Computation of general inner-outer and spectral factorization problems. IEEE Trans Autom Control 45:2307–2324
6. Weiss M (1994) Spectral and factorizations through the constrained Riccati equation. IEEE Trans Autom Control AC-26: 677–681
7. Cristian O (2005) Constructive solutions to spectral and inner-outer factorizations with respect to the disk. Automatic 41:1855–1866
8. Varga A (1998) Computation of inner-outer and spectral factorizations of rational matrices. IEEE Trans Autom Control 43(5):684–688
9. Van Dooren P (1979) The computation of Kronecker's canonical form of a singular pencil. Linear Algebra Appl 27:103–141

Author Biography

Rongwu Xiang Asscc Professor, mainly re-search about Data Analysis and Pattern Recognition

Multi-Objective Evolutionary Algorithm Based on Arena Principle and Niche

Jian-Qiu Zhang and Feng Xu

Abstract This paper researches the principle of RM-MEDA & MOEA/D, proposes Regularity Model Based Multi-Objective Estimation of Distribution Algorithm and Decomposition Algorithm. In order to solve the problem of Pareto optimal solutions, a new method with Niche Genetic Algorithm, a policy of double elite and a Pareto local search strategy. And use numerical simulation to prove the algorithm is better than NSGA-II.

Keywords Arena principle · MOEA clustering method · Niche · Pareto optimal fronts

1 Introduction

No matter in the field of natural sciences or engineering, multi-objective optimization is a very important subject and problem of research [1]. To this day, the study on multi-objective evolutionary algorithm has come to a boom period in the international community, and all kinds of the new dominated evolutionary algorithm appeared [2]. Genetic algorithm is a general optimization algorithm by the genetic and evolutionary process and the formation of creature; it is used to solve complex problems by improving the advantages [3–5]. Now the multi-objective evolutionary algorithms have been widely used in natural sciences and engineering. In the past, a wide variety of evolutionary algorithms have been used to solve multi-objective optimization problems, however, lots of several types of MOEAs, developed for problems with a low number of objectives, have difficulties to find a good Pareto Set approximation for higher dimensions as theirs high computational

J.-Q. Zhang (✉) · F. Xu
School of Science, Anhui University of Science and Technology,
Huainan, People Republic of China
e-mail: fzjq@163.com

complexity, non-elitism approach, and the need for specifying a sharing parameter [6]. In this paper, we suggest a non-dominated sorting based on Arena's principle which alleviates all the above difficulties. Specifically, the Arena non-dominated sorting approach with O (rmN) computational complexity is presented first and then a well-proportioned distribution is presented. Simulation results on tow difficult problems show that the proposed Arena Principle is able to find much better spread of solutions in all problems compared to NSGA-II. Because of Arena Principle's low computational requirements, Arena Principle should find more and more applications in the years to come [7].

2 The Evolutionary Algorithm of Arena Principle

Arena Principles are Pareto-based approaches of algorithm [8]. The population P is randomly initialized, where Nds denotes the set of selected pairs of parent solutions. In order to sort P according to level of non-domination, each solution must be compared with every other solution in the population P to find if it is dominated. With the comparison going, the non-dominated solution is selected into Nds and removed from population P. When this process is continued to find the members of the non-dominated solutions for all population members and P is null, the above procedure is repeated. The Arena Principle is outlined below:

Nonsort_base_on_areba (Rk_T)

1. Nds = ϕ, Q = P, x $\in$ Q;
2. Q = Q − x, RK = ϕ , R = ϕ;
3. While Q is not null, y $\in$ Q

   If x $\succ$ y, then Q = Q − y, go to (3);
   If y $\succ$ x, then x = y, Q = Q − y, RK = RK $\cup$ R, R = ϕ, go to (3);
   Else $R = R \cup \{y\}$, Q − Q − y, go to (3);

4. RK$'$ = {y $\in$ RK | not(x $\succ$ y)}, Nds = Nds $\cup$ {x}
5. Q = RK$'$ $\cup$ R

   If $|Q| > 1$, then go to (1);
   Else $Nds = Nds \cup Q$

Every x in current population Q should be the host one-time and compare with every other objective in the population Q. As we see, S is the number of the non-dominated objectives in population Q (the size of Q is N), comparing K times. It will get one non-dominated solution at each comparison. After (N − 1) times of comparison, k_1 dominated individuals are removed from Q. When it gets next solution, it should compare (N − k_1 − 1) times and removed k_2 dominated individuals, and so on, when getting Kth solution, it should compare (N − k_1 − k_2− $\cdots$− k_{K-1} − 1) times. In all, it should remove (k_1 + k_2+$\cdots$+ k_{K-1}) = N − s

individuals. On average, the probability of solving is the same as we get $(N - S)/K$ non-dominated solutions at each comparison. The computational complexity is:

$$T = \sum_{i=0}^{K-1} N - i - 1 - i \times (N - K)/K$$
$$= (KN - N - K)/2$$
$$= O(KN)$$

For multi-objectives evolutionary algorithms, the average complexity is O (mKN), m is the number of objectives, generally m < N. When the number of non-dominated solutions is small, the comparison times are relatively small, so the computational complexity little. The test result shows that non-dominated solutions' ratio is generally about 70 %.

3 Comparison of Optimization Results

We compare Arena Principle with niche and NSGA-II on two tests Function:

$$F_1 : \begin{cases} f_1(x) = 1 - \exp\left(-\sum_{i=1}^{n}\left(x_i - \frac{1}{\sqrt{n}}\right)^2\right) \\ f_2(x) = 1 - \exp\left(-\sum_{i=1}^{n}\left(x_i + \frac{1}{\sqrt{n}}\right)^2\right) \end{cases}$$

where $-4 \le x_1, x_2, x_3 \le 4$.

$$F_2 : \begin{cases} f_1(x) = \sum_{i=1}^{n-1}\left(-10\exp\left(-0.2\sqrt{x_i^2 + x_{i+1}^2}\right)\right) \\ f_2(x) = \sum_{i=1}^{n}\left(|x_i|^{0.8} + 5\sin(x_i)^3\right) \end{cases}$$

where $-5 \le x_1, x_2, x_3 \le 5$.

In the absence of certain restrictions on the evolution of generation, during the experiment a variable Num is added to determine whether the procedures continue, that is, when the number of non-dominated individuals is greater than Num, the procedure stops. During the experiment, some extent by coded bytes and the size of population are decided the number of non-dominated individuals. The greater the individual's coded bytes, the size of the population, the greater the number of non-dominated are individuals. Due to the impact of the above two reasons, Num should be set up in a reasonable range. If the Num is set smaller, it may not be true Pareto optimal set however, if it is set too large, it will waste plenty of computing time and affect the operation efficiency.

Table 1 Non-dominated proportion

Proportion	40 %	50 %	60 %	70 %	80 %	90 %
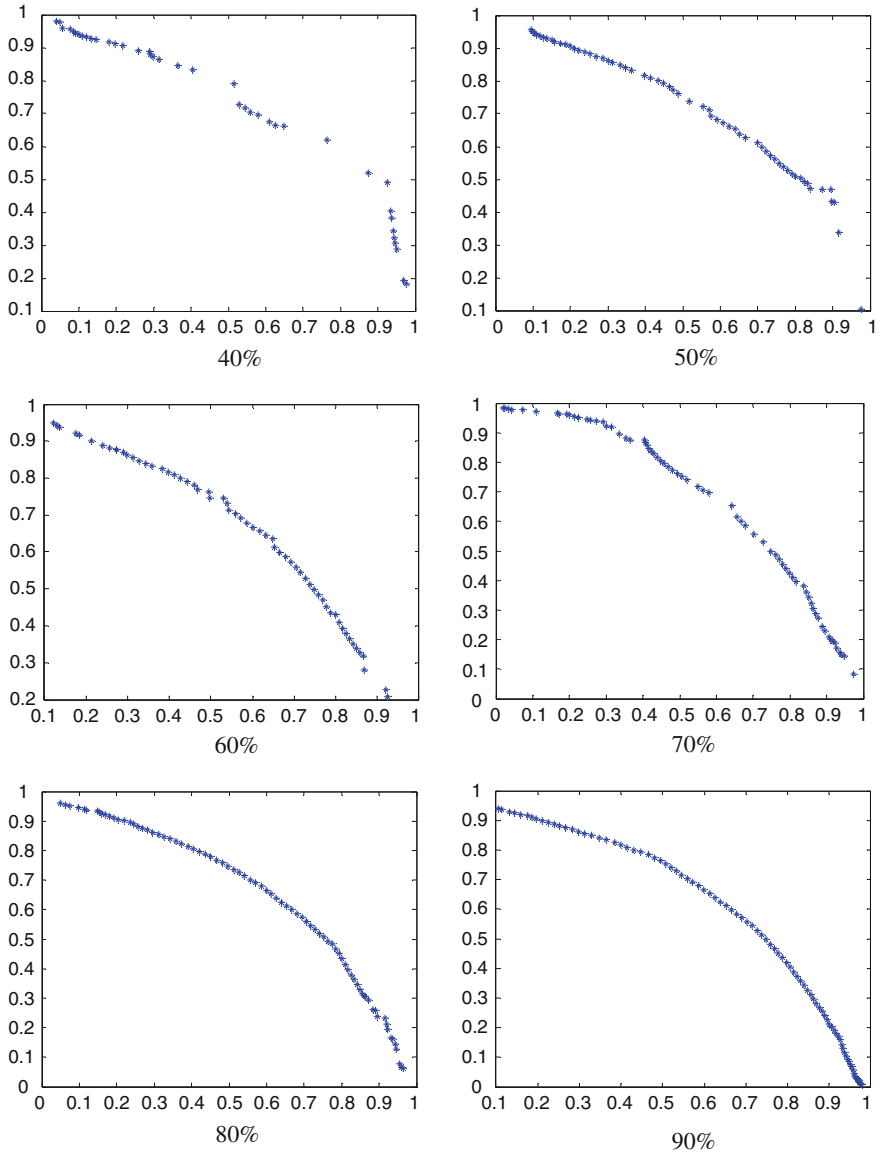
Fig. 1 Non-dominated proportion

Fig. 2 Non-dominated solutions based on arena principle with F_1
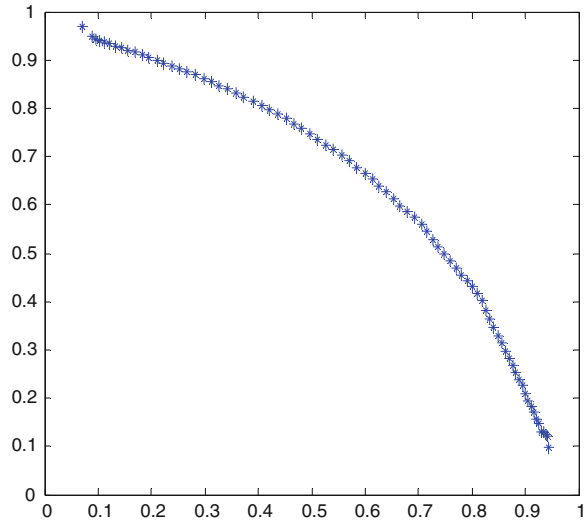
Fig. 3 Non-dominated solutions based on NSGA-II F_1
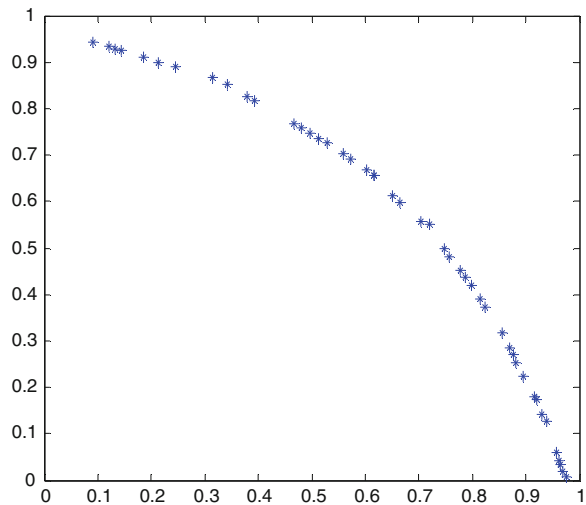
3.1 The Proportion of Non-dominated Individuals

In this test, we use the test function F_1, binary-code of individual encoding, the bytes are 8, and the size of population is 100 (Table 1).

By the following Fig. 1, it can be seen that with the number of non-dominated individuals increasing, he non-dominated individuals tend to be true Pareto optimal front. In order to get Num non-dominated individuals, it should scan and compare every individual several times in a certain population. The higher the proportion is, the more the number of times is.

Fig. 4 Non-dominated
solutions based on arena
principle with F_2
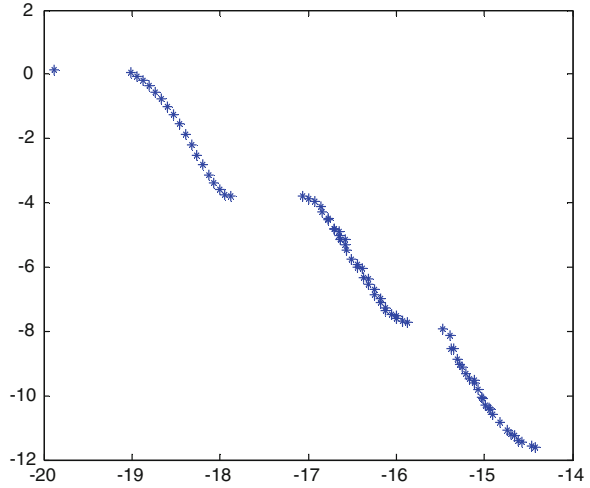
Fig. 5 Non-dominated
solutions based on NSGA-II F_2
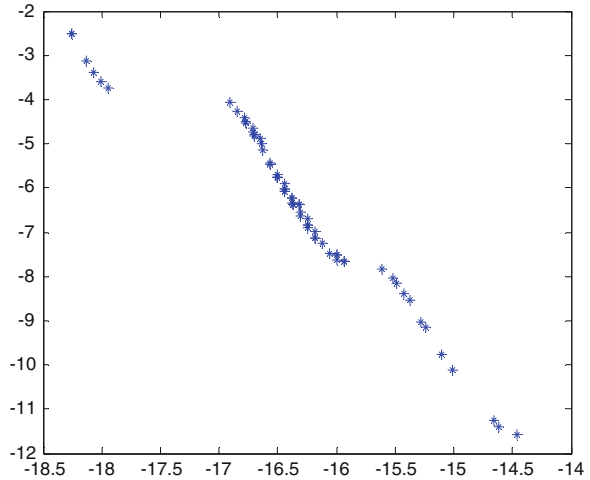
3.2 Compare with NSGA-II

For the test problems in this paper, we use the parameters as below: binary-code of
individual encoding, uniform crossover, non-uniform mutation, the number of
non-dominated individuals is 80, to NSGA-II, the evolutionary generation is 50,
objective function shared (Figs. 2, 3, 4, 5 and Tables 2, 3).

It is clear that the Arena Principle is distributing its population along the
obtained front better than *NSGA-II*. With the similar number of function evalua-
tions, Arena Principle obtains larger range of solutions and more uniform.

Table 2 Comparison of test conditions on F_1

Algorithm	Encoded type	Bytes	Size of pop	Crossover probability	Mutation probability
Arena principle	Binary	8	100	0.8	0.01
NSGA-II	Binary	8	200	0.8	0.01

Table 3 Comparison of test conditions on F_2

Algorithm	Encoded type	Bytes	Size of pop	Crossover probability	Mutation probability
Arena principle	Binary	8	300	0.8	0.01
NSGA-II	Binary	8	300	0.8	0.01

4 Conclusion

A computational fast multi-objective evolutionary algorithm based on non-dominated sorting approach is proposed in this paper. The process of construction of non-dominated set directly affects the running time and efficiency of the algorithm. In this paper, the non-dominated set based on Arena Principle is constructed and its correctness is verified through the tests. It can be seen that it reduces the number of comparisons effectively, improves the computational efficiency, and maintains the diversity of the population well uniformly.

References

1. Coello CAC, Pulido GT (2001) A micro-genetic algorithm for multi-objective optimization. In: Proceedings of first international conference on evolutionary multi-criterion optimization, doi:10.1007/3-540-44719-99
2. Horn J, Nafpliotis N, Goldberg DE (1994) A niched pareto genetic algorithm for multi-objective optimization. In: Proceedings of 1st IEEE conference evolutionary computation, pp 82–87
3. Deb K, Pratap A, Agrawal S, Meyrivan T (2002) A fast and elitist multi-objective genetic algorithm: NSGA-II. IEEE Trans Evol Comput 6:182–197
4. Deb K, Agrawal S, Pratap A, Meyrivan T (2000) A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: MSGA-II. KanGAL Report, 200001, Kanpur, Indian Institute of Technology
5. Laumanns M, Thiele L, Deb K, Zitzler E (2002) Combining convergence and diversity in evolutionary multi-objective optimization. Evol Comput 10:263–282
6. Erickson M, Mayer A, Horn J (2001) The niched Pareto genetic algorithm 2 applied to the design of groundwater remediation system. In: Zitzler E, Deb K, Thiele L, Coello CAC, Corne D (eds) Proceedings of the 1st international conference on evolutionary multi-criterion optimization, EMO 2001. Springer, Berlin, pp 681–695
7. Zitzler E, Deb K, Thiele L (1999) Comparison of multiobjective evolutionary algorithms: empirical results. Evol Comput 8:173–196
8. Knowles JD, Corne DW (2000) Approximating the non-dominated front using the Pareto archived evolution strategy. Evol Comput 8(2):149–172

Solving the Set Cover Problem in the Tile Assembly Model

Zhou Xu, Zhou Yan Tao and Li Ken Li

Abstract The tile assembly model is a novel biological computing model that is scalable and has highly parallel computing ability. In this paper, the tile assembly model can be applied to solve the set cover problem which is a well-known NP-complete problem. In order to achieve this, we design a MinSetCover system which is composed of three parts, the initial configuration subsystem, the non-deterministic choosing subsystem and the detecting subsystem. Then we analysis the system's complexity and certify the system's correctness by experiment simulation.

Keywords Np-complete problem · Minimum set cover problem · The tile assembly model

Z. Xu · Z. Y. Tao (✉) · L. K. Li
School of Information Science and Engineering, Hunan University,
Changsha, 410082 Hunan, China
e-mail: Yantao_z@hnu.edu.cn

Z. Xu
e-mail: zhouxu2006@126.com

L. K. Li
e-mail: lkl@hnu.edu.cn

Z. Xu · Z. Y. Tao
College of Electrical and Information Engineering, Hunan University,
Changsha, 410082 Hunan, China

Z. Xu
College of Mathematics and Information Engineering, Jiaxing University,
Jiaxing, 314001 Zhejiang, China

1 Introduction

Based on the mathematical concept of Wang tiles, Winfree first proposed the approach of computation by self-assembly tiles and also developed a universal tile assembly systems [1, 2]. In the tile assembly model, it first uses the DNA tile's sticky-ends to store information and then processes the information by the DNA tile's self assembly operation. Many biological systems in self-assembly tile model for the NP-complete problems are developed [3–9].

Minimum set cover problem is one of the NP-complete problems. There has been some DNA computing algorithm for it been proposed in the articles [8] and [9]. Because of the model's limitations, the proposed algorithms have the disadvantages of high error rate, operating difficultly and human intervention.

In this paper, we propose a novel MinSetCover system for the Minimum set cover problem. The MinSetCover system is composed of three parts, which are the initial configuration subsystem, the nondeterministic choosing subsystem and the detecting subsystem. Then we analysis the complexity of the proposed system and do some experimental work to shown its correctness and efficiency. Our system can not only solve the minimum set cover problem but also present clear evidence.

2 The Tile Assembly Model

The tile assembly model is an extension of a model proposed by Wang [1, 2], and was fully defined by Rothemund and Winfree. The definitions here are similar to the one in Refs. [3–6].

A tile system S can be denoted as a 6-triple $< \sum, \tau, g, T, Seed, Final >$, where T is a finite set of tiles containing empty, $Seed$ is a seed configuration, $Final$ is a final configuration, g is a strength function and $\tau \in \mathbb{N}$ is the temperature.

1. Σ: a finite alphabet representing the DAE molecules' binding domains and $null \in \Sigma$. The 4-triple $<\sigma_N, \sigma_E, \sigma_S, \sigma_W > \in \Sigma^4$ represents the sticky ends' set of a tile. For convenient, we denote an empty tile which is $<null, null, null, null >$ and it will not binding with other tiles.
2. A strength function $g: \Sigma \times \Sigma \to R$, where g is commutative and $\forall \sigma \in \Sigma$, $g(null, \sigma) = 0$, denotes the strength of the binding domains. If σ and $\overline{\sigma}$ are complement, $g(\sigma, \overline{\sigma}) = 1$.
3. τ: is the temperature. If for all $\sigma, g(\sigma, \sigma) = 1$ and $\tau = 2$ then a tile t can attach only at positions with matching binding domains on the tiles in at least two adjacent positions.
4. If C is a configuration, a tile t can attach to C at position (x, y) and gain a new configuration C' if:

$$\begin{cases} (x,y) \notin C \\ \sum_{d \in D} g(bd_d(t), bd_{d^{-1}}(A_C(d(s,y)))) \geq \tau \\ (u,v) \in Z^2, (u,v) \neq (x,y) \Rightarrow A'_C(u,v) = A_C(u,v) \\ A'_C(x,y) = t \end{cases}$$

3 Solving the Minimum Set Cover Problem in the Tile Assembly Model

3.1 Definition of the Minimum Set Cover Problem

Assume S is a finite set, $S = \{s_1, s_2, ..., s_p\}$ and $s_{\hat{p}}$ is an element in S for $1 \leq \hat{p} \leq p$. Suppose that a collection C is $\{C_1, C_2, ..., C_q\}$, where $C_{\hat{q}} \subseteq S$ for $1 \leq \hat{q} \leq q$. We denote that $|S| = p$ and $|C| = q$.

Minimum Set Cover Problem: A set cover for S is a sub-collection $C_{sub} \subseteq C$ and $C_{sub} = \{C'_1, C'_2, ..., C'_m\}$ where $|C_{sub}| = m$. Assume that $\cup C'_i = S$, C_{sub} is a set cover for S. The minimum set cover problem is to find a minimum-size set cover for S.

3.2 Solving the Minimum Set Cover Problem in the Tile Assembly Model

3.2.1 The Seed Configuration Subsystem

In the seed configuration subsystem, the inputs are encoded as the L-configuration, in which the elements of the set S are encoded in the 0th row.

Theorem 1 *Let* $\sum_{seed} = \{s_{\hat{p}}, \hat{q}, =, |\}$, T_{seed} *be the set of tiles over* $\sum_{seed}$ *as shown in* Fig. 1, $g_{seed} = 1$ *and* $\tau_{seed} = 2$. *Then the seed configuration subsystem Seed ConfSystem* $= <T_{seed}, g_{seed}, \tau_{seed}>$ *generates all the seed configurations.*

Proof Let $\sum_{seed} = \{s_{\hat{p}}, \hat{q}, =, |\}$, $T_{seed} = \{<s_{\hat{p}}, null, =, \Rightarrow, <|, null, =, null>, <|, |, \hat{q}, null>, <null, C, =, \Rightarrow, <|, null, null, \Rightarrow, <null, |, =, null>\}$ be the set of tiles shown in Fig. 1, $g_{seed} = 1, \tau_{seed} = 2$ and *SeedConfSystem* $=$
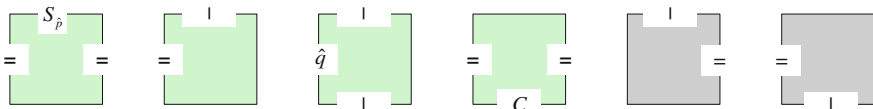
Fig. 1 The six types of tile in T_{seed}

$<T_{seed}, g_{seed}, \tau_{seed}>$. As shown in Fig. 1, there are six types of tiles in *Seed-ConfSystem* with the south and east sides as the input sides, the north and west sides as the output sides. Let C_{seed} be a seed configuration so there are some position $(x, y) \in Z^2$ such that: $\forall(x = 0, y = 0)$, $C_{seed}(x, y) = <|, null, =, null>$; $C_{seed}(x, y) = <|, null, =, null>$; $\forall(x = -1, y = 0)$, $C_{seed}(x, y) = <C, null, =, \Rightarrow$; $\forall(-p - 1 \leq x \leq -2, y = 0)$, $C_{seed}(x, y) = <s_{\hat{p}}, null, =, \Rightarrow$ where $1 \leq \hat{p} \leq p$; $\forall(x = -p - 2, y = 0)$, $C_{seed}(x, y) = <|, null, null, \Rightarrow$; $\forall(x = 0, \ 1 \leq y \leq q)$, $C_{seed}(x, y) = <|, |, \hat{q}, null>$; $\forall(x = 0, y = q + 1)$, $C_{seed}(x, y) = <null|, =, null>$; For all the other positions $(x, y) \in \mathbb{Z}^2, (x, y) \notin C_{seed}$.

3.2.2 The Nondeterministic Choosing Subsystem

Theorem 2 Let $\sum_{Nondeter} = \{C, C_{\hat{q}}^1, C_{\hat{q}}^0, \hat{q}, =\}$, $T_{Nondeter}$ be the set of tiles over $\sum_{Nondeter}$ as shown in Fig. 2, $g_{Nondeter} = 1$ and $\tau_{Nondeter} = 2$. Then the subsystem $NondeterChooseSystem = <T_{Nondeter}, g_{Nondeter}, \tau_{Nondeter}>$ generate the solution space of the set-cover problem not certainly and produce all the initial configurations.

Proof Let $T_{Nondeter} = \{<C, C, C_{\hat{q}}^1, \hat{q}>, <C, C, C_{\hat{q}}^0, \hat{q}>, <null, C, =, \Rightarrow\}$ be the set of tiles shown in Fig. 2, $g_{Nondeter} = 1$ and $\tau_{Nondeter} = 2$, and *NondeterChoose System* = $<T_{Nondeter}, g_{Nondeter} = 1, \tau_{Nondeter}>$. As shown in Fig. 2, there are three kinds of tile in *NondeterChooseSystem* with the south and east sides as the input sides, the north and west sides as the output sides. Let $C_{Nondeter}$ be an initial configuration so there are some position $(x, y) \in Z^2$ such that: $\forall(x = -1, 1 \leq y \leq q)$, $C_{NonDeter}(x, y) = <C, C, C_{\hat{q}}^0, \hat{q}>$ or $<C, C, C_{\hat{q}}^1, \hat{q}>$; $\forall(x = -1, y = q + 1)$, $C_{NonDeter}(x, y) = <null, C, =, \Rightarrow$.

3.2.3 The Detecting subsystem

After generating the initial configurations, it is very important to design a detect subsystem for identify the legal configurations. In this section, the detecting subsystem is designed to identify and check the legal solutions.

Lemma 3 Let $\sum_{Detect} = \{s_{\hat{p}}, C_{\hat{q}}^1, C_{\hat{q}}^0, OK, |, =\}$, T_{Detect} be the set of tiles over $\sum_{Detect}$ as shown in Fig. 3, $g_{Detect} = 1$ and $\tau_{Detect} = 2$. Then the detecting

Fig. 2 There are three types of tile in $T_{Nondeter}$
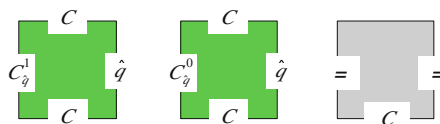
subsystem DetectSystem $= <T_{Detect}, g_{Detect}, \tau_{Detect}>$ *identifies the legal configurations and produce the final configurations.*

Proof Let $\sum_{Detect} = \{s_{\hat{p}}, C_{\hat{q}}^0, C_{\hat{q}}^1 OK, |, =\}, T_{Detect} = \{<s_{\hat{p}}, s_{\hat{p}}, C_{\hat{q}}^1, C_{\hat{q}}^1>, <s_{\hat{p}}, s_{\hat{p}}C_{\hat{q}}^0, C_{\hat{q}}^0>, <OK, s_{\hat{p}}, C_{\hat{q}}^1 C_{\hat{q}}^1>, <OK, OK, C_{\hat{q}}^1, C_{\hat{q}}^1>, <OK, OK, C_{\hat{q}}^0, C_{\hat{q}}^0>, <null, |, =, null>, <|, |, null, C_{\hat{q}}^1>, <|, |, null, C_{\hat{q}}^0>, <null, OK, =, =>\}$ be the set of tiles shown in Fig. 3, $g_{Detect} = 1, \tau_{Detect} = 2$, and *DetectSystem* $= <T_{Detect}, g_{Detect}, \tau_{Detect}>$. Let C_{Detect} be a legal final configuration so there are some $(x, y) \in Z^2$ such that: $\forall (x = -p - 2, 1 \leq y \leq q)$, $C_{Detect}(x, y) = <|, |, null, C_{\hat{q}}^1>$ or $<|, |, null, C_{\hat{q}}^0>$ where $1 \leq \hat{q} \leq q; \forall (-p - 1 \leq x \leq -1, 1 \leq y \leq q)$, $C_{Detect}(x, y) = \{<s_{\hat{p}}, s_{\hat{p}}, C_{\hat{q}}^1 C_{\hat{q}}^1>, <s_{\hat{p}}, s_{\hat{p}}, C_{\hat{q}}^0, C_{\hat{q}}^0>, <OK, s_{\hat{p}}, C_{\hat{q}}^1, C_{\hat{q}}^1>, <OK, OK, C_{\hat{q}}^1, C_{\hat{q}}^1>, <OK, OK, C_{\hat{q}}^0, C_{\hat{q}}^0>\}$ where $1 \leq \hat{p} \leq p$ and $1 \leq \hat{q} \leq q; \forall (-p - 1 \leq x \leq -1, y = q + 1)$, $C_{Detect}(x, y) = <null, OK, =, \Rightarrow>; \forall (x = -p - 2, y = q + 1), C_{Detect}(x, y) = <null, |, =, null>$ with the value '***SUC***'; For all the other positions $(x, y) \in \mathbb{Z}^2, (x, y) \notin C_{Detect}$.

3.3 The Complexity Analysis of the Proposed MinSetCover System

Lemma 4 *The assembly time of the proposed set cover system is* $O(q)$ *and the space complexity is* $O(pq)$ *and* $O(pq)$ *distinct tiles .*

Proof For the final configurations that contain the tile $<null, |, =, null>$ with the value '*SUC*', it is the complete $(p + 3) \times (q + 2)$ rectangle. The complete rectangle's depth is $q + 2$, so the assembly time is $O(q)$. The area of the complete rectangle is $(p + 3) \times (q + 2) = pq + 2p + 3q + 6$ and the space complexity is $O(pq)$.

In the initial configuration subsystem, it needs $p + q + 4$ distinct tiles. In the nondeterministic choosing subsystem, it needs $2q + 1$ distinct tiles. In the
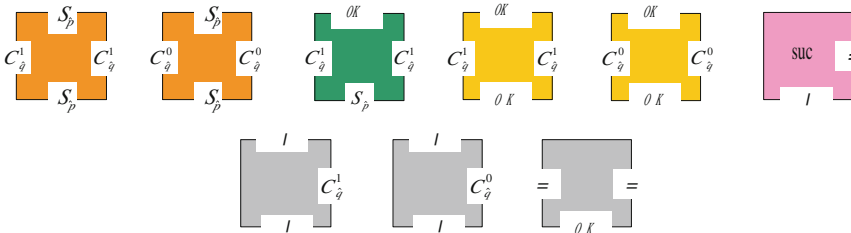
Fig. 3 There are nine kinds of tile in T_{Detect}

detecting subsystem, it needs $2pq + 4q + 2$ distinct tiles. Hence, the proposed set cover system needs $(p + q + 4) + (2q + 1) + (2pq + 4q + 2) = 2pq + 7q + p + 7$ tiles. In the proposed set cover system, it uses $O(pq)$ distinct tiles.

Lemma 5 *The probability that a single nondeterministic execution of the proposed set cover system success in attaching a* SUC *tile is at least* 0.5^q.

Proof Only the tiles in column-1 attach nondeterministically, and at each of those positions there are exactly two choices of tile t, which are $<C, C, C_{\hat{q}}^0, \hat{q}>$ and $<C, C, C_{\hat{q}}^1, \hat{q}>$. So the probability of the correct tile attaching at each location is 0.5 and there are exactly q places where such tiles can attach. Thus if there exists an answer, at least 0.5^q of the assemblies attach a *SUC* tile.

4 Conclusion

In this paper, we have designed a MinSetCover system in the tile assembly model. The MinSetCover system uses $O(pq)$ distinct tiles with $O(q)$ assembly time, $O(pq)$ space complexity and the least successive rate 0.5^q where p is equal to $|S|$ and q is equal to $|C|$. Lastly, Experimental results verify our system.

References

1. Winfree E (1998) Algorithmic self-assembly of DNA. PhD Thesis, California Institute of Technology, Pasadena
2. Winfree E (1998) Design and self-assembly of two-dimensional DNA crystals. Nature 394:1223–1226
3. Brun Y (2006) Arithmetic computation in the tile assembly model: addition and multiplication. Theor Comput Sci 378:17–31
4. Brun Y (2008) Solving NP-complete problems in the tile assembly model. Theor Comput Sci 395:31–46
5. Brun Y (2008) Solving satisfiability in the tile assembly model with a constant-size tile set. J Algorithms 63:151–166
6. Cheng Z, Huang YF (ed) (2009) Algorithm of solving the subset-product problem based on DNA Tile self-assembly. J Comput Theor Nanosc 6: 1161–1169
7. Cui G, Li C (ed) (2009) Application of DNA self-assembly on maximum clique problem. Advances in intelligent and soft computing, vol 116, pp 359–368
8. Liu J, Yang L, Li KL (ed) (2008) An $O(1.414^n)$ volume molecular solutions for the exact cover problem on DNA-based supercomputing. J Inf Comput Sci 5:153–162
9. Chang WL, Guo M (2003) Solving the set-cover problem and the problem of exact cover by 3-sets in the Adleman–Lipton model. Biosystem 72:263–275

Boundedness for Multilinear Commutators of Calderón-Zygmund Operator on Herz-Hardy Spaces

Jiangfeng Hao, Hailian Wang and Rulong Xie

Abstract In this paper, the boundedness for the multilinear commutators generated by Calderón-Zygmund operator with Lipschitz function is discussed, and obtain that it is bounded on Herz-Hardy spaces.

1 Introduction

In 2002, Pérez and Trujillo-González (see [1]) introduced a kind of multilinear commutators of singular integral operators and obtained that the sharp weighted estimates for this kind of multilinear commutators. From then, the properties of multilinear commutators have been widely studied in harmonic analysis; see [2–8]. Chen and Ma (see [5]) established the boundedness of multilinear commutators from $L^p(R^n)$ to Triebel-Lizorkin space. Mo and Lu (see [6]) obtained the bounded properties for multilinear commutators of Marcinkiewicz integrals. In this paper, we establish that the multilinear commutators of Calderón-Zygmund operator is bounded on Herz-Hardy spaces. Let us firstly recall some notation and definitions.

For $\beta > 0$, the homogeneous Lipschitz space $\dot{\wedge}_\beta(R^n)$ is the space of function f such that

---

J. Hao · H. Wang · R. Xie (✉)
Department of Mathematics, Chaohu University, Chaohu 238000, China
e-mail: rulongxie@163.com

J. Hao
e-mail: starhjf@163.com

H. Wang
School of Mathematical Sciences, Anhui University, Heifei 230039, China
e-mail: wanghailian77@163.com

$$||f||_{\dot{\Lambda}_\beta} = \sup_{x,h \in R^n; h \neq 0} |\triangle_h^{[\beta]+1} f(x)|/|h|^\beta < \infty.$$

where $\triangle_h^k$ denotes the kth difference operator (see [9]).

Given any positive integer m, for $1 \leq i \leq m$, we denote by C_i^m the family of all finite subsets $\sigma = \{\sigma(1), \sigma(2), \ldots, \sigma(i)\}$ of $\{1, 2, \ldots, m\}$ with i different elements. For any $\sigma \in C_i^m$, the complementary sequences σ' is given by $\sigma' = \{1, 2, \ldots, k\} \backslash \sigma$. Let $\mathbf{b} = (b_1, b_2, \ldots, b_m)$ be a finite family of locally integrable functions. For all $1 \leq i \leq m$ and $\sigma = \{\sigma(1), \ldots, \sigma(i)\} \in C_i^m$, we denote $\mathbf{b}_\sigma = (b_{\sigma(1)}, \ldots, b_{\sigma(i)})$ and the product $b_\sigma = b_{\sigma(1)} \ldots b_{\sigma(i)}$. With this denote, if $\beta_{\sigma(1)} + \cdots + \beta_{\sigma(i)} = \beta_\sigma$, we write

$$||\mathbf{b}_\sigma||_{\dot{\Lambda}_{\beta_\sigma}} = ||b_{\sigma(1)}||_{\dot{\Lambda}_{\beta_{\sigma(1)}}} \cdots ||b_{\sigma(i)}||_{\dot{\Lambda}_{\beta_{\sigma(i)}}}.$$

For the product of all the function, we simply write $||\mathbf{b}||_{\dot{\Lambda}_\beta} = \prod_{i=1}^m ||b_i||_{\dot{\Lambda}_{\beta_i}}$, where $\sum_{i=1}^m \beta_i = \beta$.

Definition 1   A kernel $K(.,.) \in L^1_{loc}(R^n \times R^n \backslash \{(x,y) : x = y\})$ is called a Calderón-Zygmund kernel if

$$|K(x,y)| \leq C|x - y|^{-n}, \tag{1}$$

for $x \neq y$; and if $2|y - z| \leq |x - z|$,

$$|K(x,y) - K(x,z)| + |K(y,x) - K(z,x)| \leq C \frac{|y - z|}{|x - z|^{n+1}}. \tag{2}$$

The Calderón-Zygmund operator associated with the above kernel $k(.,.)$ is defined by

$$Tf(x) = p.v. \int_{R^n} K(x,y)f(y)dy. \tag{3}$$

We defined the multilinear commutators of Calderón-Zygmund operators with Lipschitz function as follows:

$$T_{\mathbf{b}}(f)(x) = \int_{R^n} \prod_{i=1}^m (b_i(x) - b_i(y))K(x,y)f(y)dy. \tag{4}$$

where $\mathbf{b} = (b_1, b_2, \ldots, b_m)$ be a finite family of locally integrable functions with $b_i \in \dot{\Lambda}_\beta(R^n)$, $1 \leq i \leq m$.

Let $B_k = \{x \in R^n : |x| \leq 2^k\}$, $C_k = B_k \backslash B_{k-1}$, and $\chi_k = \chi_{C_k}$ for $k \in Z$, where χ_{C_k} is the character function of set C_k.

Definition 2   (see [10]) Let $\alpha \in R$, $0 < p < \infty$, $0 < q < \infty$.

1. The homogeneous Herz space $\dot{K}_q^{\alpha,p}(R^n)$ is defined by

$$\dot{K}_q^{\alpha,p}(R^n) = \{f \in L_{loc}^q(R^n \setminus \{0\}) : ||f||_{\dot{K}_q^{\alpha,p}(R^n)} < \infty\},$$

where

$$||f||_{\dot{K}_q^{\alpha,p}(R^n)} = \{\sum_{k=-\infty}^{\infty} ||f\chi_k||_{L^q}^p\}^{\frac{1}{p}};$$

2. The nonhomogeneous Herz space $K_q^{\alpha,p}(R^n)$ is defined by

$$K_q^{\alpha,p}(R^n) = \{f \in L_{loc}^q(R^n) : ||f||_{K_q^{\alpha,p}(R^n)} < \infty\},$$

where

$$||f||_{K_q^{\alpha,p}(R^n)} = \{||f\chi_{B_0}||_{L^q}^p + \sum_{k=1}^{\infty} ||f\chi_k||_{L^q}^p\}^{\frac{1}{p}}.$$

Definition 3 (see [10]) Let $\alpha \in R$, $0 < p \leq \infty$, $0 < q \leq \infty$.

1. The homogeneous Herz-Hardy space $H\dot{K}_q^{\alpha,p}(R^n)$ is defined by

$$H\dot{K}_q^{\alpha,p}(R^n) = \{f \in S'(R^n) : G(f) \in \dot{K}_q^{\alpha,p}(R^n)\}.$$

Moreover, $||f||_{H\dot{K}_q^{\alpha,p}(R^n)} = ||G(f)||_{\dot{K}_q^{\alpha,p}(R^n)}$.
2. The nonhomogeneous Herz-Hardy space $HK_q^{\alpha,p}(R^n)$ is defined by

$$HK_q^{\alpha,p}(R^n) = \{f \in S'(R^n) : G(f) \in K_q^{\alpha,p}(R^n)\}.$$

Moreover, $||f||_{HK_q^{\alpha,p}(R^n)} = ||G(f)||_{K_q^{\alpha,p}(R^n)}$, where $G(f)$ is the grand maximal function of f.

The Herz-Hardy space have the atomic decomposition characterization.

Definition 4 (see [11]) Let **b** be as above such that $b_i \in \dot{\wedge}_{\beta_i}(R^n)$, $1 \leq i \leq m$, $0 < \beta_i < 1$, $\sum_{i=1}^m \beta_i = \beta < n$, $1 < q < \infty$ and $\alpha \geq n(1 - \frac{1}{q})$. A function $a(x)$ is said to be a center $(\alpha, q, \mathbf{b})$ atom (or a center $(\alpha, q, \mathbf{b})$ atom of restrictive type), if it satisfies the following condition:

1. supp $a \subset B(0, l) = \{x \in R^n : |x| \leq l\}$ for some $l > 0$ (or for some $l \geq 1$);
2. $||a||_{L^q} \leq |B(0, l)|^{-\frac{\alpha}{n}}$;
3. $\int_{R^n} a(x) x^{\mu} dx = \int_{R^n} a(x) x^{\mu} \prod_{j \in \sigma} b_j(x) dx = 0$, $|\mu| \leq s$, $s \geq [\alpha + n(\frac{1}{q} - 1)]$ for any

$\sigma = (\sigma_1, \ldots, \sigma_i) \in C_m^i$, $i = 1, \ldots, m$.

A temperate distribution f is said to belong to the $H\dot{K}_{q,\mathbf{b}}^{\alpha,p}(R^n)$ (or $HK_{q,\mathbf{b}}^{\alpha,p}(R^n)$) , if it can be written as $f = \sum_j \lambda_j a_j$ in the S' sense, where a_j is a center $(\alpha, q, \mathbf{b})$ atom (or a center $(\alpha, q, \boldsymbol{b})$ atom of restrictive type), $\lambda_j \in C$ and $\sum_j |\lambda_j|^p < \infty$.

Moreover, we define the quasinorm on $H\dot{K}_{q,\mathbf{b}}^{\alpha,p}(R^n)$ (or $HK_{q,\mathbf{b}}^{\alpha,p}(R^n)$) by

$$||f||_{H\dot{K}_{q,\mathbf{b}}^{\alpha,p}(R^n)}\big(\text{ or } ||f||_{HK_{q,\mathbf{b}}^{\alpha,p}(R^n)}\big) = \inf\left\{ (\sum_j |\lambda_j|^p)^{1/p}\right\},$$

for all decomposition of $f = \sum_j \lambda_j a_j$.

Obviously, $H\dot{K}_{q,\mathbf{b}}^{\alpha,p}(R^n)$ (or $HK_{q,\mathbf{b}}^{\alpha,p}(R^n)$) is a subspace of $H\dot{K}_q^{\alpha,p}(R^n)$(or $HK_q^{\alpha,p}(R^n)$).

2 Main Results and Proofs

Theorem 2.1 *Suppose that $T_{\mathbf{b}}(f)$ is defined by* (4), *where $\mathbf{b}$ is as above such that* $b_i \in \dot{\wedge}_{\beta_i}(R^n), 1 \leq i \leq m, 0 < \beta_i, \sum_{i=1}^m \beta_i = \beta < 1, 0 < p < \infty, 1 < q_1, q_2 < \infty, \frac{1}{q_1} - \frac{1}{q_2}$ $= \frac{\beta}{n}$ *and* $n(1 - \frac{1}{q_1}) \leq \alpha < n(1 - \frac{1}{q_1}) + \beta$. *Then $T_{\mathbf{b}}(f)$ is bounded from $H\dot{K}_{q_1,\mathbf{b}}^{\alpha,p}(R^n)$ to* $\dot{K}_{q_2}^{\alpha,p}(R^n)$.

Remark 2.1 Theorem 2.1 is also hold for the nonhomogeneous Herz-Hardy space.

In order to proving the main results, we need the following Lemmas.

Lemma 2.1 *(see* [12]*) Let $0 < \alpha < n$, $1 < p < \frac{n}{\alpha}$ and $\frac{1}{q} = \frac{1}{p} - \frac{\alpha}{n}$. Then the fractional integral operator T_α defined by*

$$T_\alpha(f)(x) = \int_{R^n} \frac{f(y)}{|x - y|^{n-\alpha}} \, dy$$

is bounded from $L^p(R^n)$ to $L^q(R^n)$.

Lemma 2.2 *Suppose that $T_{\mathbf{b}}(f)$ is defined by* (4), *where $\mathbf{b}$ is as above such that* $b_i \in \dot{\wedge}_{\beta_i}(R^n)$, $1 \leq i \leq m$, $0 < \beta_i < 1$, $\sum_{i=1}^m \beta_i = \beta < n$, $1 < p < \frac{n}{\beta}$ *and* $\frac{1}{q} = \frac{1}{p} - \frac{\beta}{n}$. *Then $T_{\mathbf{b}}(f)$ is bounded from $L^p(R^n)$ to $L^q(R^n)$.*

Proof As $b_i \in \dot{\wedge}_{\beta_i}(R^n)$ $(0 < \beta_i < 1)$ for any $i = 1, 2, \ldots, m$. Then

$$|T_{\boldsymbol{b}}(f)(x)| \leq \int_{R^n} |\prod_{i=1}^m (b_i(x) - b_i(y)) K(x, y) f(y)| dy$$

$$\leq \int_{R^n} |\prod_{i=1}^m (b_i(x) - b_i(y))||f(y)| \frac{1}{|x - y|^n} dy$$

$$\leq C||\mathbf{b}||_{\dot{\wedge}_\beta} \int_{R^n} \frac{|f(y)|}{|x - y|^{n-\beta}} dy.$$

Note that $0 < \sum_{i=1}^{m} \beta_i = \beta < n$, by Lemma 2.1, therefore

$$||T_{\mathbf{b}}(f)||_{L^q} \leq C||\mathbf{b}||_{\dot{\Lambda}_\beta}||T_\beta(|f|)||_{L^q} \leq C||\mathbf{b}||_{\dot{\Lambda}_\beta}||f||_{L^p};$$

Thus we finish the proof of Lemma 2.2.

Proof of Theorem 2.1. Let $f \in H\dot{K}_{q_1,\mathbf{b}}^{\alpha,p}(R^n)$ and $f(x) = \sum_{j=-\infty}^{\infty} \lambda_j a_j(x)$ be the atomic decomposition for f as in Definition 4. We write

$$||T_{\mathbf{b}}(f)||_{\dot{K}_{q_2}^{\alpha,p}(R^n)}^p \leq C \sum_{k=-\infty}^{\infty} 2^{k\alpha p} \left(\sum_{j=-\infty}^{k-3} |\lambda_j| ||T_{\mathbf{b}}(a_j)\chi_k||_{L^{q_2}} \right)^p$$

$$+ C \sum_{k=-\infty}^{\infty} 2^{k\alpha p} \left(\sum_{j=k-2}^{\infty} |\lambda_j| ||T_{\mathbf{b}}(a_j)\chi_k||_{L^{q_2}} \right)^p := I_1 + I_2.$$

For I_2, by the Lemma 2.2, we obtain that (L^{q_1}, L^{q_2}) boundedness of $T_{\mathbf{b}}$.

When $0 < p \leq 1$, we have

$$I_2 \leq C \sum_{k=-\infty}^{\infty} 2^k \left(\sum_{j=k-2}^{\infty} |\lambda_j| ||a_j\chi_k||_{L^{q_2}} \right)^p$$

$$\leq C \sum_{j=-\infty}^{\infty} |\lambda_j|^p \left(\sum_{k=-\infty}^{j+2} 2^{(k-j)\alpha p} \right)$$

$$\leq C \sum_{j=-\infty}^{\infty} |\lambda_j|^p \leq C||f||_{H\dot{K}_{q_1,\mathbf{b}}^{\alpha,p}(R^n)}^p.$$

When $p > 1$, we have

$$I_2 \leq C \sum_{k=-\infty}^{\infty} 2^{k\alpha p} \left(\sum_{j=k-2}^{\infty} |\lambda_j| ||a_j\chi_k||_{L^{q_2}} \right)^p$$

$$\leq C \sum_{j=-\infty}^{\infty} |\lambda_j|^p \left(\sum_{k=-\infty}^{j+2} 2^{(k-j)\alpha p/2} \right) \left(\sum_{k=-\infty}^{j+2} 2^{(k-j)\alpha p'/2} \right)^{p/p'}$$

$$\leq C \sum_{j=-\infty}^{\infty} |\lambda_j|^p \leq C||f||_{H\dot{K}_{q_1,\mathbf{b}}^{\alpha,p}(R^n)}^p.$$

For I_1, note that $x \in C_k$, $y \in C_j$, $j \leq k-3$, we have $|x-y| \sim |x|$. We consider

$$||T_{\mathbf{b}}(a_j)\chi_k||_{L^{q_2}} = \left\{ \int_{C_k} \left[| \int_{C_j} \prod_{i=1}^{m} (b_i(x) - b_i(y)) K(x,y) a_j(y) dy | \right]^{q_2} dx \right\}^{1/q_2}$$

$$\leq C \{ \int_{(C_k} [\int_{C_j} \prod_{i=1}^{m} |b_i(x) - b_i(0)| |K(x,y) - K(x,0)| |a_j(y)| dy]^{q_2} dx \}^{1/q_2}$$

$$+ C \{ \int_{C_k} [\int_{C_j} \sum_{i=1}^{m-1} \sum_{\sigma \in C_i^m} |(b(x) - b(0))_\sigma| |(b(y) - b(0))_{\sigma'}| |K(x,y)$$

$$- K(x,0)| |a_j(y)| dy]^{q_2} dx \}^{1/q_2}$$

$$+ C \{ \int_{C_k} [\int_{C_j} \prod_{i=1}^{m} |b_i(y) - b_i(0)| |K(x,y)| |a_j(y)| dy]^{q_2/2} dx \}^{1/q_2}$$

$$:= II_1 + II_2 + II_3.$$

By the Minkowski's inequality and Hölder's inequality, we have

$$II_1 = C \{ \int_{C_k} [\int_{C_j} \prod_{i=1}^{m} |b_i(x) - b_i(0)| K(x,y) - K(x,0)| |a_j(y)| dy]^{q_2} dx \}^{1/q_2}$$

$$\leq C \{ \int_{C_k} [\int_{C_j} \prod_{i=1}^{m} |b_i(x) - b_i(0)| \frac{|y|}{|x|^{n+1}} |a_j(y)| dy]^{q_2} dx \}^{1/q_2}$$

$$\leq C \int_{C_j} \{ \int_{C_k} (\prod_{i=1}^{m} |b_i(x) - b_i(0)| \frac{|y|}{|x|^{n+1}})^{q_2} dx \}^{1/q_2} |a_j(y)| dy$$

$$\leq C \int_{C_j} \{ \int_{C_k} (||\mathbf{b}||_{\dot{\Lambda}_\beta} |x|^\beta| \frac{|y|}{|x|^{n+1}})^{q_2} dx \}^{1/q_2} |a_j(y)| dy$$

$$\leq C ||\mathbf{b}||_{\dot{\Lambda}_\beta} 2^{k\beta} 2^{-kn} 2^{(j-k)} 2^{kn/q_2} 2^{-j\alpha} 2^{jn(1-\frac{1}{q_1})}$$

$$\leq C ||\mathbf{b}||_{\dot{\Lambda}_\beta} 2^{-j\alpha} 2^{(j-k)[n(1-\frac{1}{q_1})+1]}.$$

For II_2, similar to estimate II_1, using the Minkowski's inequality and Hölder's inequality, we obtain

$$II_2 \leq C \sum_{i=1}^{m-1} \sum_{\sigma \in C_i^m} \int_{C_j} \{ \int_{C_k} [|(b(x) - b(0))_\sigma| |(b(y) - b(0))_{\sigma'}| \frac{|y|}{|x|^{n+1}}]^{q_2} dx \}^{1/q_2} |a_j(y)| dy$$

$$\leq C \sum_{i=1}^{m-1} \sum_{\sigma \in C_i^m} \int_{C_j} \{ \int_{C_k} (||\mathbf{b}_\sigma||_{\dot{\Lambda}_{\beta_\sigma}} |x|^{\beta_\sigma} ||\mathbf{b}_{\sigma'}||_{\dot{\Lambda}_{\beta_{\sigma'}}} |y|^{\beta_{\sigma'}} \frac{|y|}{|x|^{n+1}})^{q_2} dx \}\}^{1/q_2} |a_j(y)| dy$$

$$\leq C ||\mathbf{b}||_{\dot{\Lambda}_\beta} 2^{-j\alpha} 2^{(j-k)[n(1-\frac{1}{q_1})+\beta_{\sigma'}+1]}.$$

Let us turn to estimate II_3. By the Minkowski's inequality and Hölder's inequality, we have

$$II_3 \leq C\{\int_{C_k}[\int_{C_j}\prod_{i=1}^{m}|b_i(y)-b_i(0)|K(x,y)||a_j(y)|dy]^{q_2}dx\}^{1/q_2}$$

$$\leq C\int_{C_j}\{\int_{C_k}(\prod_{i=1}^{m}|b_i(y)-b_i(0)||x-y|^{-n})^{q_2}dx\}^{1/q_2}|a_j(y)|dy$$

$$\leq C\int_{C_j}\{\int_{C_k}(||\mathbf{b}||_{\dot{\Lambda}_\beta}|y|^\beta|x|^{-n})^{q_2}dx\}^{1/q_2}|a_j(y)|dy$$

$$\leq C||\mathbf{b}||_{\dot{\Lambda}_\beta}2^{-j\alpha}2^{(j-k)[n(1-\frac{1}{q_1})+\beta]}.$$

Therefore, when $0 < p \leq 1$, we get

$$I_1 \leq C||\mathbf{b}||_{\dot{\Lambda}_\beta}^p\sum_{k=-\infty}^{\infty}2^{k\alpha p}\{\sum_{j=-\infty}^{k-3}|\lambda_j|(2^{-j\alpha}2^{(j-k)[n(1-\frac{1}{q_1})+1]}+2^{-j\alpha}2^{(j-k)[n(1-\frac{1}{q_1})+\beta]})\}^p$$

$$\leq C\sum_{j=-\infty}^{\infty}|\lambda_j|^p\{\sum_{k=j+3}^{\infty}(2^{(j-k)[n(1-\frac{1}{q_1})+1-\alpha]}+2^{(j-k)[n(1-\frac{1}{q_1})+\beta-\alpha]})^p\}$$

$$\leq C\sum_{j=-\infty}^{\infty}|\lambda_j|^p \leq C||f||_{H\dot{K}_{q_1,\mathbf{b}}^{\alpha,p}(R^n)}^p.$$

When $p > 1$, we get

$$I_1 \leq C||\mathbf{b}||_{\dot{\Lambda}_\beta}^p\sum_{k=-\infty}^{\infty}2^{k\alpha p}\{\sum_{j=-\infty}^{k-3}|\lambda_j|(2^{-j\alpha}2^{(j-k)[n(1-\frac{1}{q_1})+1]}+2^{-j\alpha}2^{(j-k)[n(1-\frac{1}{q_1})+\beta]})\}^p$$

$$\leq C\sum_{j=-\infty}^{\infty}|\lambda_j|^p\{\sum_{k=j+3}^{\infty}(2^{(j-k)[n(1-\frac{1}{q_1})+1-\alpha]}$$

$$+2^{(j-k)[n(1-\frac{1}{q_1})+\beta-\alpha]})^{p/2}\}\{\sum_{k=j+3}^{\infty}(2^{(j-k)[n(1-\frac{1}{q_1})+1-\alpha]}+2^{(j-k)[n(1-\frac{1}{q_1})+\beta-\alpha]})^{p'/2}\}^{p/p'}$$

$$\leq C\sum_{j=-\infty}^{\infty}|\lambda_j|^p \leq C||f||_{H\dot{K}_{q_1,\mathbf{b}}^{\alpha,p}(R^n)}^p.$$

Thus the proof of Theorem 2.1 is completed.

References

1. Perez C, Trujillo-Gonza R (2002) Sharp weighted estimates for multilinear commutators. J London Math Soc 65:672–692
2. Gorosito O, Pradolini G, Salinas O (2007) Weighted weak-type estimates for multilinear commutators of fractional integrals on spaces of homogeneous type. Acta Math Sin (English Series) 23:1813–1826
3. Hu GE, Meng Y, Yang DC (2005) Multilinear commutators of singular integrals with non doubling measures. Integr Equ Oper Theorey 51:235–255
4. Meng Y, Yang DC (2006) Multilinear commutators of Calderón-Zygmund operators on Hardy-type spaces with non-doubling measures. J Math Anal Appl 317:228–244
5. Chen YP, Ma BL (2007) The Bounedeness of Multilinear Commutators form $L^p(R^n)$ to Triebel-Lizorkin space. Anal Theor Appl 23:112–128
6. Mo HX, Lu SZ (2007) Boundedness of generalized higher commutators of Marcinkiewicz integrals. Acta Math Sci 27B:852–866
7. Xie RL, Shu LS (2008) On multilinear commutators of θ- type Calderón-Zygmund operators. Anal Theor Appl 24:260–270
8. Lian JL, Li J, Wu HX (2011) Multilinear commutators of BMO functions and multilinear singular integrals with non-smooth kernels. Appl Math J Chin Univ 26:109–120
9. Paluszynski M (1995) Characterization of the Besov spaces via the commutator operator of Coifman, Rochberg and Weiss. Indiana Univ Math J 44:1–17
10. Liu LZ, Lu SZ (2006) Continuity for maximal multilinear Bochner-Riesz operators on Hardy and Herz-Hardy spaces. Acta Math Sin (English Series) 22:69–76
11. Lu SZ, Yang DC (1995) The weighted Herz-type Hardy spaces and its applications. Sci China (Series A) 38:662–673
12. Stein EM (1993) Harmonic analysis: real variable methods. Princeton University Press, Princeton NJ, Orthogonality and Oscillatory integrals

Algorithm of DNA Computing Model for Gate Assignment Problem

Zhixiang Yin, Min Chen and Qingyan Li

Abstract In the core of airport operation, aircraft stands assignment (ASA) is a typical kind of combinatorial optimization. In this paper, by analyzing the ASA problem, gate assignment problem is transferred to vertex coloring model. A DNA computing model for airport gate assignment is proposed. The simulation results show that the algorithm compared with other optimization is very easy and feasible.

Keywords Gate assignment · Vertex coloring · DNA computing

1 Introduction

Airport gate assignment [1] is an important decision making problem which involves multiple and conflict objectives in airport. In certain extent, command center of airport assign boarding gate to scheduled flight, confirm flights are operating normally, serve passengers boarding gate. Airport gate assignment problem mainly solves two problems: time and the space of flights occupy. It has relationship with the kinds of scheduled flights, arrival and departure intensive degree, equipment of airport, and flight distribution method. The main purpose is assigning flights or aircraft to appropriate terminal and ramp positions (gates) and maintaining safety, security, cost efficiency and conveniences to passengers.

Z. Yin (✉) · M. Chen · Q. Li
Anhui University Of Science and Technology, Huainan, Anhui, People's Republic of China
e-mail: zxyin66@163.com

M. Chen
e-mail: chenmin0925@163.com

Q. Li
e-mail: 13625620162@163.com

The key to solve this question is to establish apportion model of airport gate assignment and optimize. The task is a real-world decision problem which involves multiple and conflicting objective while respecting various constraints, and a number of mathematical models and analytical tools have been developed by researchers in recent years. There are two main optimization objectives, one is expert system [2, 3], allocation principle based on establishing knowledge base system, consider more non-quantitative criterion; the other is mathematical programming, minimizing the total walking distance, using 0–1 integer programming to discuss the feasibility of assignment and how to assign. Flight delays are quite universal in real-life, and the initial assignments would be disrupted easily even with unexpected changes of flight schedules [4–8]. Models have been proposed to achieve the even distributions of gate idle times by minimizing dispersion of idle time periods for static aircraft-gate assignment problem. Xu created mathematical model based on minimum time of passenger to transfer, adopted tabu search (TS) algorithm [9]. Ding considering the limit resource of airport, improve Xu's algorithm [10]. Li [11] solve the question by the matching between time and space, establish mathematics model, use tabu search (TS) algorithm to optimize. Tian Chen and others propose and implements a strategy based on genetic algorithm to solve the problem, the algorithm has advantages including computation time, real time performance and optimization over other existed algorithms [12].

The paper proposes and implements a strategy based on genetic algorithm to solve the problem. In order to solve problem of flight delays, paper requires method of airport gate assignment has strong robustness. Considering adding a certain buffer time between the continuous flights at the same flight stand can help solve this question. This paper design an algorithm based on model of DNA computing on graph vertex coloring, systematically analyzes situation of domestic airport gate position during busy period and management of airport operating, considering matching between flights type and flight stand.

2 Graph Vertex Coloring Problem

The field of DNA computing in its present form started with Professor Adleman's [13] celebrated experiment in 1994. He solved Hamilton path on DNA coding, which attracted many scholars to pay attention to DNA computing. Biochemical reaction theory based DNA computation is of the massive inherent parallelism, so compared to silicon computer, DNA computer has most superiority out and away on NP problems. Recently great progress both in theory and biological operation technology has been made in DNA computing research, more research achievements are obtained in most fields. Many scholars has built models of DNA computing, used these models to solve NP-complete problem in Graph Theory and Optimizations [14–17].

The graph coloring problem (GCP) is a widely studied combinatorial optimization problem with good application background, which is widely used in many

fields, such as Process problems, Time-table problems, Storage problem and so on. Anyway, the graph coloring problem is NP-complete problem.

Let $G = (V, E)$ be a finite, simple and undirected graph with the set of vertices $V(G)$ and the set of edges $E(G)$, and $C = \{1, 2, \ldots k\}$, a set of k colors. Given an undirected graph, the vertex coloring problem is to assign a different color for vertex mutually adjacent. When the vertices v_i and v_j have different color, it's said there is an edge between them. For a positive integer k, f is a mapping from $V(G)| \to C(k), f: V(G)| \to C(k), f(v_i) \neq f(v_j), \forall e = v_i v_j \in E(G)$. The graph coloring problem is find the minimum number k, if k is fond, the graph G is called k—vertices coloring.

3 Aircraft Stands Assignment

3.1 Airport Gate Assignment Problem

According to the scale of plane, aircraft Stand is divided into three parts: large, middle, small, and each plane should be assigned to a suitable parking place in the limited time from the beginning of parking to the plane leaving. In different periods of a day, it is necessary to affirm the kind of aircraft stand to be parked and time, satisfy's the following constraint condition

1. Each scheduled flight should be distributed to one stand.
2. At the same time and stand, only a scheduled flight is assigned.
3. Satisfy requirements: the space from beginning to end of a flight should not be shorter than minimum time of an airplane to finish its work (containing loading and unloading of passenger and cargo, cleaning airplane and so on), furthermore, time interval is necessary to two or more flights to avoid the accident.
4. Gate position and space scheduled flight needs should match each other. In other words, larger flight is only assigned to larger gate position, middle gate position or small gate position is not allowed, but middle scheduled flight can be assigned to larger gate position.

3.2 Graph Vertex Coloring Model of Airport Gate Assignment

Graph vertex coloring problem of airport gate assignment is solved by graph model: $G = (V, E)$, $V = \{v_i|$ scheduled flight i waiting to be assigned$\}$; $E = \{(v_i, v_j)|$ the two scheduled flight collide$\}$. In the model of airport gate assignment, V is the set of all scheduled flights, $V = \{v_i|i = 1, 2, 3, \ldots, n\}$, E is the set of scheduled flight which is in conflict with each other, $E = \{(v_i, v_j)|i=1, 2, \ldots, m, j = 1, 2, \ldots n, i \neq j\}$, so it is expressed as function $f: V(G)| \to C(K); f(v_i) \neq f(v_j), \forall e = v_i v_j \in E(G)$, all of the following criteria must be met:

1. At the same time and stand, only a scheduled flight is assigned.
2. Each scheduled flight should be distributed to one stand. Since it begins to work, the stand is not permitted to be assigned other scheduled flight.
3. Gate position and space scheduled flight needs should match each other.

4 Algorithm Design of Airport Gate Assignment Problem

4.1 The Definition of Time Slice

The judgment method of the time conflict of the scheduled flights is the key to the question. Time slice concept is introduced to solve this question. At a certain stage, time is divided into several parts, each time period is occupied by several flights which are not assigned the same stand at the same time because of the time conflict of the scheduled flights, is called time slice. By introducing definition of time slice, a certain stage of time is divided into several time slices. The start/stop moment of scheduled flight occupied can also be expressed one or more continuous time slices. At the same time slice, t only allows one flight, one aircraft stand.

4.2 Algorithm Design

Considering the reality of the situation of airport gate assignment, a algorithm for the graph vertex coloring problem of airport gate assignment has been presented, follow the steps outlined below:

Step 1 According to partitioning algorithm of time slice, set E of scheduled flights which have conflict with others in time is determined.

Step 2 Binary map $G = (V, E)$ is given by time conflict situation.

Step 3 Coding each vertex. Using prepared protein probe to confirm different series of colors of each vertex. Based on existing coding theory, determine the mathematics constraint conditions by constraint conditions of biological experiment, and code, get link sequence.

Step 4 Build data pool. Protein probe represented as each vertex and DNA sequence represented as different colors are built together in right ways according to some certain synthetic method. DNA chain is built to represented as possible solution of problem, obtain the input data by DNA computing.

Step 5 Search data set which meets collision condition, and read out the results of the calculation.

Step 6 DNA sequencing are employed to identify operation result, using PCR, find the position of the connection sequence, and determine the solution of graph coloring.

4.3 DNA Algorithms and Implementations

4.3.1 Biological Preparation

Probe is a bit of strand or small interfering RNA (about 20–500 bp), prepared for realization of DNA biological algorithm, made use of testing complimentary sequence of nucleic acid. Protein probe (take antibody for example) only reacts with target molecules by ion and hydrogen bond on a few specific parts. On the other hand, nucleic acid probe reacts with its complementary strand on more than ten, hundred, thousand parts according to the different length of hybrid molecule. Because nucleic acid has annealing characteristics, nucleic acid probe is the only probe used to DNA computing. Two nucleic acid probes are prepared to make to choose and delete in this algorithms: Vertex probe (P_v) and Color probe (P_M).

The vertex coding must reflects neighbouring relations and vertex coloring. Coding of vertex v is divided into three parts, denoted as PMQ. M represents its corresponding color k $(2 \leq k \leq n)$, both P and Q are constructed by two parts. P is constructed by prefix P_P and suffix N_p of the order N of vertex v; Q is constructed by back end N_S of order of vertex v and suffix Q_S, $P = P_P N_p, Q = N_S Q_S$. Coding of vertex v can be expressed as $PMQ = P_P N_p M N_S Q_S$. If degree of a vertex v is less then $k - 1 (k \geq 2)$, some corresponding vertex should be deleted in order to reduce the solution space (data pool). In process of choose vertexes, these vertexes are not allowed to be chosen in order to avoid the second delete operation. A kind of corresponding probe P_v is made to take out the DNA coding of vertex we need in the test tube, Fig. 1 is the form and interlinking of probe P_v. Each vertex needs $n - 1$ probes $P_{v1}, P_{v2}, \ldots, P_{v(n-1)}$ which respectively corresponds to $n - 1$colors.

If adjacent vertex v_i and v_j has the same color M, connect their corresponding DNA chain $P_{pi} N_{pi} M N_{si} Q_{si}$ and $P_{pj} N_{pj} M N_{sj} Q_{sj} = \overline{Q_s} N_{pj} M N_{sj} Q_{sj}$ to create DNA long chain, the long chain which doesn't meet coloring requirement of DNA chain should be deleted. In order to finish delete operation, each adjacent vertex create v_i, v_j and each kind of color M need to make their corresponding probe P_M, Fig. 2 is the form and interlinking of probe $P_M (P_{pj} = \overline{Q_s})$. X is a DNA chain, and $L(N_{si} Q_{si} M N_{sj} N_{sj}) > L(Y) > L(N_{si} Q_{si} N_{pj})$, the distance from Y to $N_{si} Q_{si} N_{pj}$ is long enough to ensure that the two don't connect with each other. Then, probe P_M, vertex v_i and v_j with the same color M, DNA chain only links with complementary base, form the structure of pre-hairpin. It requires $L(N_{si} Q_{si} M N_{sj} N_{sj}) > L(Y)$, DNA long chain built by three vertexes, probe P_M doesn't delete it (the length is not enough) when the first and the third vertex has the same color (they are not adjacent because of the coding vertex).

Fig. 1 Probe P_v
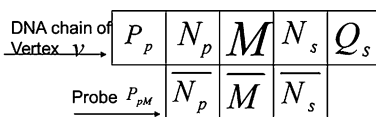
Fig. 2 Probe P_M
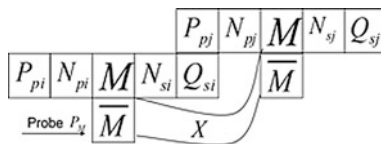
4.3.2 Biochemical Process for Implementing

In order not to confuse matters, all coding (DNA segment) *PMQ* about vertex *V* in test tube is denoted as {v}, using the set of vertexes represents the collection of vertex, at the same time, reflects adjacent relation and all coloring coding.

Step 1 Create the coding of all vertexes. The coding about vertex reflect adjacent relation and *n* kinds coding of colors *PMQ*, put them in initial test tube. Copy it to be standby.

Step 2 k = 2.

Step 3 Put probe $P_{vi}, P_{v2}, \ldots, P_{vk}(v \in V_1 \cup V_2 \cup \cdots \cup V_{q+1})$ to tube T_0, take out all vertexes whose color coding is 1 or 2, put to tube T_1, melting, to confirm the tube T_1 doesn't contains $k + 1, \ldots n - 1$ coding to vertexes $V_1 \cup V_2 \cup \cdots \cup V_{q+1}$.

Step 4 Build data pool.

Step 4.1 Put ligase to tube T_1, and react, create DNA chain.

Step 4.2 Delete DNA chain does not meet the criteria. According to construct of P_M, DNA long chain created by all adjacent vertexes with the same time can be deleted, the left part is put to tube T_2. Test tube T_2 is the solution space (data pool).

Step 5 Search all satisfied vertex coding, and outputs a calculation result. According to the definite of coding, it's known to all graph *G* is *k*—vertexes coloring. There exists a DNA chain containing all vertexes $V_1 \cup V_2 \cup \cdots \cup V_{q+1}$ from v_{11} through *W*. It's the longest chain in data pool, the lengths is $D(T_1) \times L(PMQ) - D(D(T_1) - 1) \times L(P_P) = (5D(T_1) + 1) \times L(P_P)$. So, if there exists a chain in data pool T_2, the graph is *k*—vertexes coloring.

Step 5.1 Add buffered solution to test tube, use gel electrophoresis to DNA chain.

Step 5.2 Get DNA chain of maximum molecular weight, the chain is the output result.

Step 6 DNA sequencing to result. Delete gel in DNA chain, read base sequence of the left chain.

Step 7 When the graph is vertexes coloring, the number of color is $X(G)$, else, $k = k + 1$, to Step 3 (using ready T_0 to create new data pool T_1).

Step 8 End.

5 Conclusion

The biggest characteristic of DNA computing is large-scale parallel computing and the huge information storage capacity. Based on airport gate assignment problem, the paper presents algorithm of graph vertex coloring, use DNA computing model

to solve airport gate assignment problem successfully, get the number of vertex coloring. Examples proves the algorithm of DNA computing can achieve in theory.

References

1. Wen J, Li B, Wang QR, Du W (2005) Graph coloring model and algorithm of gate assignment airport. Syst Eng-Theor Methodol Appl 14(2):136–140
2. Gosling GD (1990) Design of an expert system for aircraft gate assignment. Transp Res A 24(1):59–69
3. Su YY, Srihari KA (1993) Knowledge based aircraft gate assignment advisor. Comput Ind Eng 25(2):123–126
4. Xu J, Bailey G (2001) The airport gate assignment problem: mathematical model and a tabu search algorithm. In: Proceeding of the 34th Hawaii international conference on system sciences IEEE, pp 102–111
5. Ding H, Lim A, Rodrigues B et al (2004) Aircraft and gate scheduling optimization at airports. In: Proceeding of the 37th international conference on system sciences IEEE, pp 74–81
6. Li W, Liu C, Tu F (2006) Optimized assignment of civil airport gate. J Nanjing Univ Acronautics and Astronautics 38(4):433–437
7. Chen T, Guixi X (2005) Airport parking-position assignments strategy based on GA. Comput Eng 31(3):186–188
8. Adleman LM (1994) Molecular computation of solution to combinational problem. Science 266(5187):1021–1024
9. Ouyang Q, Kaplan PD, Liu S et al (1997) DNA solution of the maximal clique problem. Science 278(5337):446–449
10. Liu Q, Wang L, Frutos AG et al (2000) DNA computing on surfaces. Nature 403(6766):175–179
11. Braich RS, Chelyapov N, Johnson C et al (2000) Solution of a 20-variable 3-SAT problem on a DNA computer. Science 296:499–502
12. Yin ZX, Pan LQ, Shi XL (2006) DNA algorithm of minimal spanning tree. In: Proceedings of SPIE, vol 6358. pp 1–5
13. Braaksma JP, Shortreed JH (1971) Improving airport gate usage with critical path. Transp Eng ASCE 97(2):187–203
14. Babic O, Teddorovic D, Tosic V (1984) Aircraft stand assignment to minimize walking. J Transp Eng 110(1):55–66
15. Mangoubi DFX, Mathaisel RS (1985) Optimizing gate assignments at airport terminals. Transp Sci 19:173–188
16. Chang G, Wei SM (2006) Study of model based on combinatorial optimization for aircraft stands assignment. J Civ Aviat Univ China 24(3):28–31
17. Yan S, Huo C (2001) Optimization of multiple objective gate assignment. Transp Res 35A:413–432

The Advantages and Disadvantages of DNA Password in the Contrast to the Traditional Cryptography and Quantum Cryptography

Jun Jiang and Zhixiang Yin

Abstract In recent years, DNA Password is a new cryptography field which appears with DNA calculation. This paper briefly introduces the DNA cryptography, Traditional cryptography and Quantum cryptography. From the basic concepts, theoretical foundations, concrete operation processes, safety basis and the current development achievements and shortcomings of the three fields, this paper obtains DNA cryptography's advantages and disadvantages compared with traditional DNA cryptography and quantum cryptography and analyzes its causes. Then, this paper shows the achievements of DNA cryptography in practical application and its future development direction. Finally, it will make an outlook that a new mixed cryptography system can be created on the advantages of the three kinds of cryptography.

Keywords DNA computing · DNA cryptography · Traditional cryptography · Quantum cryptography

1 Introduction

DNA calculation is a frontier disciplines in recent years [1–3], while cryptography is a nearly two-thousand-year old tool used to protect data from the old Caesar password to modern cryptography overdue. However, the connections between the two seemingly no contact disciplines are closer in modern technology.

J. Jiang · Z. Yin (✉)
Anhui University of Science and Technology, Huainan 232001, China
e-mail: zxyin66@163.com

J. Jiang
e-mail: jiangjun_for_work@163.com

Cryptography technology is the core of the information security technology [4], comprising mainly the password encoding technology and cryptanalysis [5, 6]. Currently, the classification standard for password theory and technology is whether it is based on mathematical problem. One of the cryptography not based on mathematical problem is DNA password, which will be discussed in this paper.

The encryption, decryption and security of traditional cryptography and modern cryptography, which based on the former, are all based on math problems. All the password systems but the one-time ones are theoretical secure. If the attacker has enough computing power, he could decipher these passwords. However, as the most highly developed and frequently used password systems, they have quiet mature algorithm and theory.

As a new conception, Quantum cryptography uses the theoretical basis as its safe mode [6], simply, it is a new password system based on single photon and its inherent quantum properties [7]. Quantum state can't be determined without interfering the system, namely, the attacker can't decode the password without being discovered. And according to Heisenberg's Uncertainty Principle, quantum password will be unbreakable.

DNA password is a new kind of cryptography associated with DNA computing research in recent years. Using DNA as its information carrier, and modern biotechnology its implementation tool, DNA password makes use of DNA's advantages to realize its function of encryption. With the rapid development of DNA computing and the mature of biotechnology, the rising DNA cryptography is also developing in theory and practice. It is widely believed that DNA cryptography, Traditional cryptography and Quantum cryptography will become three branches of cryptography [8–10]. However, as the advantages and disadvantages of DNA password are too prominent [11], it will take a long time to build a theoretical framework and practical operating system which can keep pace with its two forerunners [12, 13]. At present, the academic research into DNA password is gradually developing [14, 15].

2 Modern Cryptography and Quantum Cryptography

2.1 Conception of Cryptography

The basic work manner of an encryption system is called cryptosystem. The two basic elements of a cryptosystem are cryptographic algorithm and secret key.

A cryptosystem usually has five parts: plaintext space U; cryptogram space L; key space K. $k = <K_e, K_d>$ (K_e is encryption key, K_d is decryption key); encryption algorithm E; decipherment algorithm D. Plain text $m \in U$, secret key $k = <K_e, K_d> \in K$, encryption $C = E_{K_e}(m) \in L$, decryption $m = E_{K_d}(C) \in U$.

2.2 Modern Cryptography

Modern cryptography is developed on the basis of classical cryptography or traditional cryptography. Traditional cryptography bases on various mathematical problems such as Non-Polynomial Time Complete (NP-C), and has already established a relatively complete theoretical system. Using traditional password is most convenient. Computer, DNA computer even quantum computer could be used to calculate; Wire, optical fiber, radio channel and so on could be used to transport; Optical disk, magnetic medium, DNA and any other storage could be used to keep information. It is easy to implement public key encryption, private key encryption, authentication and digital signature and many other functions. It is the advantages of traditional cryptography, but there are insurmountable difficulties in actual use.

Traditional password but the one-time ones is theoretical secure. That is, if the attacker has infinite computing power, only one-time password could be used in traditional password. However, one-time password is restricted in actual use because it is very hard to keep a cipher.

2.3 Quantum Cryptography

2.3.1 Conception of Quantum Cryptography

The new unbreakable password developed from physics is quantum cryptography. It uses quantum as information carrier and transport by quantum channel. Different from traditional cryptography, quantum cryptography is a new conception, which relies on physics theory, especially quantum physics as its safe model. In short, it is a system based on single photon and their inherent quantum properties.

2.3.2 Implementation Process of Quantum Cryptography

Roughly, the process goes like this: send the encrypted information on the ground, then the information goes through the atmosphere and is received by the satellite; finally, the satellite transmits the information to receiving stations. So far, there has been three main kinds of implementation processes:

1. Uncertainty principle;
2. Bell principle;
3. Non-orthogonal quantum states.

2.3.3 Advantages of Quantum Cryptography

The security of quantum cryptography is built on the basis of quantum physics. Even if the attacker has a strong computing power, he still could not break the decoding-proof Quantum Cryptography. The laws of physics assure the safety of quantum channel. Even if the attacker can do the things he want to, and has unlimited computing power, even P = NP, he could not decipher quantum password. At present, optical fiber is used to transmit numerous data rapidly and steadily. So, how to develop and use an easier and faster way has become the reason and purpose of quantum cryptography research.

The security of quantum cryptography is built on Heisenberg's uncertainty principle, principle of quantum no-cloning and single photon indivisibility. To comply with the laws of physics can guarantee unconditional security. According to uncertainty principle, certain pairs of quantities can never be measured simultaneously with arbitrarily high precision. If the data transmitted by photon is intercepted by an attacker, he can't decipher the password without being discovered, thus, ensure the security of the password.

2.3.4 Problems of Quantum Cryptography

Seen from the above, there are some problems to be solved;

1. Low temperature state: quantum encryption needs low temperature;
2. Encryption speed;
3. Reception: quanta might scatter in the atmosphere, so it is hard for them to be received by certain satellite.

3 DNA Passwords

3.1 Conception of DNA Cryptography

DNA password is a new kind of password. Using DNA as its information carrier, and modern biotechnology its implementation tool, DNA password makes use of DNA's high storage density, high parallelism and low consumption to realize its functions of encryption, authentication and signature. To know DNA password, one must know DNA calculation and DNA biotechnology first.

In 1994, Aldeman, using modern molecular techniques, raised the Hamiltonian Path Problem and found its DNA solution. After the successful tests, DNA calculation was widely used to solve NP problems. The basic method goes like this: Use the special double helix DNA structure and nucleotide complementary pairing principle; encode the data, mapping the object of operation into DNA molecule.

Then, use biological method to make a data pool in the tube. After that, in accordance with certain rules, we map the data operation of the original problem into controlled DNA biochemical process. Finally, get the result of the operation through molecular biotechnology such as PCR, mutagenesis, molecular structure purification, electrophoresis, immuno-magnetic beads. This process could be simplified as:

1. For a special issue, encode the problem related data into DNA chain.
2. Under appropriate conditions, mix the copies of the DNA chain in solution. According to the principle of base pairing, the short chains will automatically combine into longer ones.
3. After the reaction, filter out the results that satisfy certain conditions by PCR, immuno-magnetic beads. That is the result wanted.

After years of research, there have been several models for DNA calculation. Although these models have different working and reacting process, their algorithm models are the same in core method and logical process.

Although DNA password developed from DNA calculation, they are different. DNA calculation uses DNA technology to solve mathematical problem, while DNA password uses biology problems as its safety basis [16, 17].

3.2 Implementation Process of DNA Password

Supposing the despatcher is Alice, and she has K_A. The recipient is Bob, he has K_B (If $K_A = K_B$, it is symmetric encryption, while $K_A \neq K_B$, it is public key encryption). Alice uses K_A to change M into C by E. Unless he has K_B, it is difficult to get M from C. Here, E is the encryption process, C is ciphertext. After Bob receives C, he can use K_B and C to get M by D. Here, K_A, K_B and C are not limited to digital, they could be DNA or any other materials.

The general scheme of DNA password:

Step A   Key generation. In this step, the encrypt key and decrypt key are formed. Usually, a DNA sequence would be designed as the primer sequence for encryption and decryption, or Watson–Crick Complementary pair of probe set.

Step B   Encryption. Change plaintext into cryptogram, (1) preprocess plaintext information: Choose a coded system to change plaintext into data. Here it is DNA system. Regard DNA nucleotide as quaternary coding, and three nucleotide stand for a letter, for example, CGA represent A, CCA represent B. (2) make Message Sequence Number: add the primer of 3′ and 5′ those have 20 nucleotide to DNA message sequence made in (1).

Step C   Transmission. Break Human gene sequence into nucleotide double chain with 50–100 nucleotide. Then change into a single chain, which will be used as redundant DNA. Then confound the DNA sequence which

contains information to among them, pray formed on a carrier to colorless micro point. Then through the ordinary unclassified way to transmission.

Step D   Decryption. The receiver and sender share encodings and primers. After get transfer vector containing the message of DNA sequence, the receiver extract the DNA in micro-point, and amplify message sequence through PCR, finally get DNA sequence by sequencing. Thus, the plaintext could be restored according to the pre-agreed coding mode.

3.3 Main Problems Facing DNA Password

1. Lack theoretical support. There has been no related theory about DNA password by far. What is the implementation model? What is the safety basis? How to achieve the goal? These problems are still unsolved. Because of these, there is almost no suitable DNA program.
2. Costliness. Among the existing program, artificial message DNA sequence should be added in the process of encryption and decryption, in order to do PCR and DNA sequencing. This needs a well-equipped biological laboratory thereby limiting the use of DNA password in actual operation.

3.4 Advantages of DNA Password

DNA password comes into being after traditional password and quantum password, so it is bound to have its unique distinction and advantage.

1. DNA is very small in size. Nano scale storage can be achieved easily;
2. Large scale parallel. Both encryption and decryption can be quickened;
3. Use biological problems as the security basis of DNA password to achieve a new password system which can resist quantum attack [11, 12].
4. In actual operation, DNA password does not require too much real time, but can manage data encryption, safety data storage, identity authentication, digital signature and information hiding.
5. Because DNA Micro-points not easy to be found, even if the ciphertext is intercepted, it is not easy to sequence the unknown DNA mixture. So Message Sequence Number is not easy to be found, thus assure the security.

3.5 Disadvantages of DNA Password

There are three main facts: insufficient development of DNA calculation; flaws of DNA password; difficulties in implementation:

1. Insufficient development of DNA calculation

First, every current DNA computational model can only solve one kind of problem (mainly the NP-problems), but not have an integrated mode or universal property as computers do;

Second, among the existing program, most face the "index explosion in the solution space". With the increase of the problem scale, the amount of DNA molecule needed increased sharply. For example, if use DNA nucleotide to represent all the 380 binary numbers, the mess of nucleotide needed surpass that of the sun's. The number of DNA molecules is not unlimited, that is the reason why Dan Boneh's method can only break the symmetric cryptography following 64 digits.

Third, in some models, a lot of human power and reaction steps are needed, reducing the possibility of large scale application of DNA password;

Finally, in the current achievements we made, every step needs human reprocessing. The process can't be automated, which is the hardest restriction to the development of DNA password.

2. Flaws of DNA password

First, it cannot be used alone. As the most developed password system, traditional cryptography is unbreakable if the attacker does not have enough computing power. The safety factory can be easily improved by increasing the key length. DNA password is hard to change, equivalent to one-time pad. It is not flexible in reuse.

Second, the achievement of DNA calculation cannot be applied directly to DNA password design and research. The encryption and decryption process of DNA password can be regarded as a process of calculation, while not all DNA calculations are about cryptography.

Finally, DNA encoding has a direct impact on the accuracy and decipher difficulty of DNA password.

3. Difficulties in implementation

From the introductions above, it can clearly be seen that:

First, encrypting and decrypting DNA password should be done in well-equipped biological laboratory. It costs too much and is difficult to be widely used.

Second, the carrier and primer in DNA password can only be transported through physical ways. While coded system and biotic condition can be transmitted as a message. However, it is electronic computers but not DNA computers that are widely used, which is not convenient, but this could guaranty the use of password.

Finally, how to protect the carrier from pollution in order to achieve the accuracy of the information passed.

4 Compare DNA Password with Traditional Password and Quantum Password

DNA password, traditional password and quantum password are based on different disciplines. They have different theories, implement manner and safety basis. This paper mainly discuss the advantages and disadvantages of DNA password, so we will compare DNA password with the other two.

4.1 The Difference Between DNA Password and Traditional Password

1. Theoretical basis: traditional password is based on mathematical problems such as NP-C, while DNA password is based on modern biology;
2. Storage media: traditional password use computer chips, while DNA password use DNA chain or DNA chips;
3. Storage capacity: 1 g silicon chip can store 16 MB, 1 g DNA can store 108 TB;
4. Operator manner: traditional password use serial arithmetic, while DNA password use parallel arithmetic;
5. Security: traditional password is based on mathematical problems, while DNA password is based on mathematical and biological problems.

4.2 The Difference Between DNA Password and Quantum Password

1. Theoretical basis: quantum password is based on quantum physics, DNA password is based on modern molecular biology;
2. Transmission mode: quantum password transport information by quantum through quantum channel, while DNA password use DNA as the carrier and deliver via a physical process;
3. Security: based on uncertainty principle, quantum password is unbreakable, while DNA password take the limitations of biology techniques as its security basis.

5 Practicability of DNA Password

DNA password can't replace traditional password and become a widely used system, however, as a supplement, DNA password may try the following things:

1. Store a One-time pad for traditional password;
2. Construct troopdoor one-way function in public key encryption;

3. Succeed previously in steganography;
4. Form biological authentication code;
5. Could be a solution to Chinas large resident identity management system.

6 Development and Prospect of DNA Password

DNA password is not suitable to replace traditional password, but as a password system with extreme advantages and disadvantages, it can provide directions in theory and actual use for the successor.

For example, mix DNA password with traditional password and quantum password to form another new system which has the advantages below

1. Mature theoretical basis as traditional password or quantum password do;
2. Ultra-large-scale parallelism, high storage density, low consumption;
3. The immunity to calculation;
4. Easy operation, repeated use.

Though the new system is too ideal and inevitably to meet difficulties in theory, design and practice, even too complicated to be used in modern encrypted transmission, it may generate a new password system in the research, or perfect the systems we already have.

References

1. Adleman L (1994) Molecular computation of solutions to combinational problems. Science 266:1021–1024
2. Amo M, Paun G, Rozenbergd G (2002) Topics in the theory of DNA computing. Theoret Comput Sci 287:3–38
3. Karl L (1997) DNA computing: arrival of biological mathematics. Math Intell 19(2):9–22
4. Feng D, Pei D (1999) Cryptography guidance. Science Press (in Chinese), Beijing
5. Wang Y, Liu J (1999) Network security theory and technology. Xian university of electronic science and technology press
6. Zhang Z (2002) Quantum computing and communication encryption. Central China Normal University Press, Wuhan
7. Quantum computation and quantum information (2004) Tsinghua University Press, Beijing
8. Boneh D, Dunworth C, Lipton RJ (1996) Breaking DES using a molecular computer. DNA Based Computers. American Mathematical Society, Providence USA, pp 37–65
9. Lu M, Lai X, Xiao G et al (2007) Based on DNA technology symmetric encryption method. Sci China 37(2):175–182
10. Cui G, Qin L, Wang Y et al (2007) DNA in the calculation of the information security technology. Comput Eng Appl 43(20):139–142
11. Clelland CT, Risca V, Bancroft C (1999) Hiding messages in DNA microdots. Nature 399(6736):533–534

12. Leier A et al (2000) Cryptography with DNA binary strands. Biosystems 57(1):13–22
13. Gehani, A, LaBean T, Reif J (2004) DNA based cryptography. Aspects Mol Comput, pp 34–50
14. Xiao G et al (2006) New field of cryptography: DNA cryptography. Chin Sci Bull 51(12):1413–1420
15. Lu M et al (2007) Symmetric-key cryptosystem with DNA technology. Sci China Ser F: Inf Sci 50(3):324–333
16. Lai X et al (2010) Asymmetric encryption and signature method with DNA technology. Sci China Inf Sci 53(3):506–514
17. Leier A, Richter C, Banzhaf W et al (2000) Cryptography with DNA binary strands. Biosystems 57(1):13–22

The Principle and Algorithm for Generating Incidence Matrix for Any Arbitrary Network

Wei Zhang, Chao-bo Lu and Hai-bo Li

Abstract In this paper, a new principle and algorithm for obtaining the incidence matrix for any arbitrary network which were represented by nodes and segments while we have already known the endpoints of each line segments in 2D space were introduced. In addition, a calculated procedure was compiled by C++ language and two extra examples were calculated. The results shown that the principal and algorithm we stated were right for auto-generating of the incidence matrix for any arbitrary network.

1 Introduction

Incidence matrix based on graph theory for representing the characteristic of arbitrary network which would been simplified with nodes and line segments representing the objects under investigation [1–3] has a wide application on expressing the topology of the network [4–6]. It has been studied extensively in many application fields such as power system [7–9], social science [10, 11], iconography [12] and so on [11–15]. Previous studies mainly concentrated on the characteristic of the incidence matrix such as spectra and eigenvalues [16–20], few focused on the generating of the incidence matrix itself [21, 22]. The essence of incidence matrix implementing connectivity analysis is to reflect the direct connection relationship between nodes which represent the positions in the fracture network. However, there is no research on auto-generating of incidence matrix for arbitrary network with nodes and segments when we just know the position of the

W. Zhang (✉) · C. Lu · H. Li
Institute of Rock and Soil Mechanics, Chinese Academy of Sciences,
Wuhan, 430071 Hubei, China
e-mail: abot866@163.com

two endpoints in each line segment in 2D space. It is an key step to further analyze the path of fluid flow or diffusion in fracture network.

In this paper, the principle and algorithm will be described in details for auto-generating of the incidence matrix which represents the connection relationship of an arbitrary network. In order to vividly express the principle and algorithm, a specific example will be elucidated.

2 Topology Analysis

In graph theory, the connection relationship between nodes can be expressed by incidence matrix. If there are n nodes which include line segments intersections and endpoints in an arbitrary network, the incidence matrix C is an n-by-n square matrix. Elements c_{ij} in C is 0 or 1, which could tell us the relationship between two nodes: when node i is relevant to node j, if the i, j nodes are connected directly, the element $c_{ij} = 1$, otherwise, $c_{ij} = 0$, including $i = j$, representing the i, j nodes is unconnected directly. As Fig. 1 shown, we could get the incidence matrix C below.

$$C_{ij} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad (i \, or \, j = 1 \text{ to } 11)$$

In the matrix C, the element $c_{37} = 1$ account for directly connection between node 3 and node 7; $c_{13} = 0$ account for there is no connected between node 1 and node 3.

Fig. 1 An arbitrary network
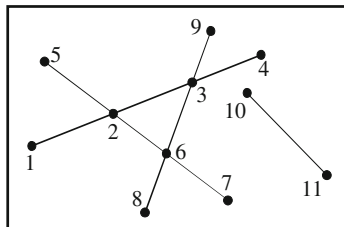
3 Principle and Algorithm

As shown in Fig. 1, one can write out the incidence matrix by hand easily for simple network with only few nodes which including endpoints of each line segment and intersections which intersected by each pair of segments. However, it will be very onerous by hand to get the incidence matrix for complex network contains large number of line segments. A new simple and feasible principle and algorithm would be presented as follow for getting the incidence matrix for any arbitrary network.

3.1 Intersections Calculated and Numbered

The network is composed of nodes representing spatial position and line segments representing channels. The two endpoints of each segment are already known and the intersections of each pair of segments will be calculated. The solved processing is shown in Fig. 2. Finally, the nodes would be reordered by coordinate of nodes increasing and labeled them with continuous serial number. The key algorithm with pseudo-code was shown as follow:

```
For L_i (i = 1 to m)
{
  For L_j (j = 1 to m)
  {
      If (L_i ∩ L_j==1)
      {
         L_i ∩ L_j = n_ij;
      }

  }
}; END
```

where m is the total number of line segments; $L_j \cap L_j == 1$ means that each pair of segments (L_i, L_j) are intersected and the intersection are located both on the line segment L_i and line segment L_j; n_{ij} is the serial number of intersection.

Similarly, we could get all the intersections on each line segment as Fig. 3 shown and reorder nodes by the nodes' position increasing (X value increasing in this paper) and renumbered the nodes. Noting that the nodes serial number on each line segment must be continuous because of the element of partitioned incidence matrix is strictly dependent on it.
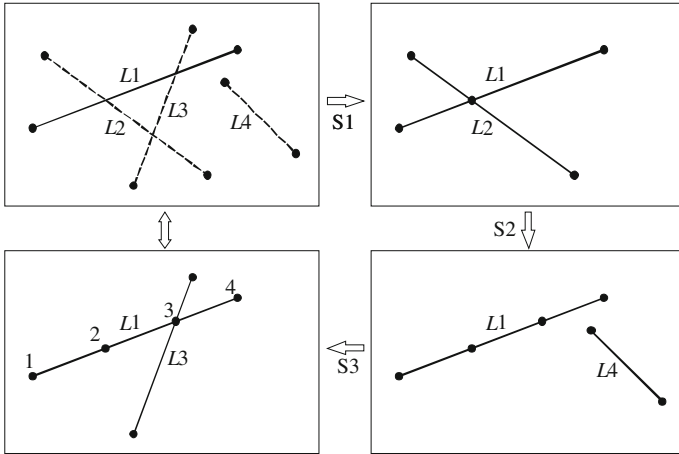
Fig. 2 The process for calculating and numbering nodes on each segment

3.2 Decompose the Network to Line Sets and Generate Partial Matrix

The network is composed of line segments and nodes on it. Firstly, we will break down the network into line segments and intersection nodes on the line segments one by one. Then we can write out the partitioned incidence matrix for the nodes on each line segment easily.

Fig. 3 Decompose network with repeating nodes
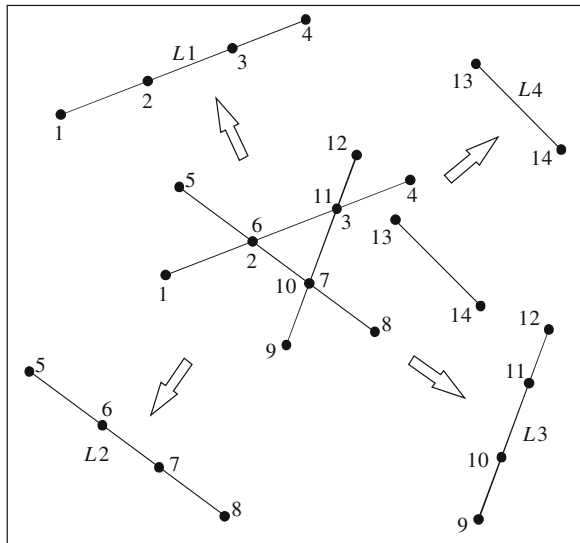
As Fig. 3 shown, the network could be decomposed to L_1, L_2, L_3 and L_4. Nodes 1, 2, 3, 4 are on line segment L_1; Nodes 5, 6, 7, 8 are on line segment L_2; Nodes 9, 10, 11, 12 are on line segment L_3 and nodes 13, 14 are on line segment L_4. There are four partitioned incidence matrices C_{L1}, C_{L2}, C_{L3} and C_{L4} for corresponding line segments. For instance, the incidence matrix C_{L1} of the nodes on line segment L_1 is

$$C_{L1ij} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (i\ or\ j = 1, 2, 3, 4)$$

Similarly, we can get that C_{L2}, C_{L3} and C_{L4} are

$$C_{L2ij} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (i\ or\ j = 5, 6, 7, 8)$$

$$C_{L3ij} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (i\ or\ j = 9, 10, 11, 12)$$

$$C_{L4ij} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad (i\ or\ j = 13, 14)$$

where i, j are the serial numbers of the nodes.

For the partitioned incidence matrix C_L of the nodes on line segments, the other elements of the matrix C_L are 0 in the total nodes matrix. The algorithm is as follow:

For L_r ($r = 1$ to m), m is the total number of the line segments in the network, and the elements in the partitioned incidence matrix C_{Lrij} for each line segment are:

$$c_{Lrij} = 0, \ \text{when } |i - j| > 1; c_{Lrij} = 1, \ \text{when } |i - j| = 1.$$

3.3 Compose the Initial Matrix for Total Nodes

After getting the partitioned incidence matrix for each line segment, we can sum them directly to be an initial matrix C_I which represents the relationship of all nodes. It allows us to get the initial incidence matrix C_{LI} as below:

$$\mathbf{C}_{LIij} = \begin{bmatrix} \mathbf{C}_{L1ij} & & & \\ & \mathbf{C}_{L2ij} & & \\ & & \mathbf{C}_{L3ij} & \\ & & & \mathbf{C}_{L4ij} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad (i \text{ or } j = 1 \text{ to } 14)$$

The algorithm for getting the initial matrix $\mathbf{C}_{LI}$ is:

$$\mathbf{C}_{LI} = \sum_{1}^{m} \oplus \mathbf{C}_{Lr} \quad (i, j = 1 \text{ to } n);$$

where n is the total number of nodes in the network.

3.4 Ranks with Repeated Nodes Operate

Next, we need to perform the logical operation of the matrix on the basis of repeated nodes, whose definition in this paper is space coincidence. If the node q and node p ($p < q$) are repeated nodes, the corresponding row q OR row p as *new* row p and the column q OR column p as *new* column p in the incidence matrix. The "OR" operation is Boolean OR operation which will be stated at the end of this section. In this paper, the nodes 2 and 6, 3 and 11, 7 and 10 are repeated nodes. Therefore, the matrix $\mathbf{C}_{LI}$ becomes $\mathbf{C}_{LA}$

$$
C_{IAij} =
\begin{bmatrix}
C_{L1ij} & & & \\
& C_{L2ij} & & \\
& & C_{L3ij} & \\
& & & C_{L4ij}
\end{bmatrix}
$$

$$
=
\begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}
\quad (i\ or\ j = 1\ to\ 14)x
$$

Because of the repeated nodes representing only one position, we should delete the repeated nodes and just keep one. The C_{IAij} would become an adjacent matrix without double nodes as follows

$$
C_{LAij} =
\begin{bmatrix}
C_{L1ij} & & & \\
& C_{L2ij} & & \\
& & C_{L3ij} & \\
& & & C_{L4ij}
\end{bmatrix}
$$

$$
=
\begin{bmatrix}
0 & 1 & 0 & 0 & 0 & x & 0 & 0 & 0 & x & x & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & x & 1 & 0 & 0 & x & x & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & 0 & x & 1 & 0 & 0 & x & x & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & x & 0 & 0 & 0 & x & x & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & x & 0 & 0 & 0 & x & x & 0 & 0 & 0 \\
x & x & x & x & x & x & x & x & x & x & x & x & x & x \\
0 & 1 & 1 & 0 & 0 & x & 0 & 1 & 1 & x & x & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & x & 1 & 0 & 0 & x & x & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & x & 1 & 0 & 0 & x & x & 0 & 0 & 0 \\
x & x & x & x & x & x & x & x & x & x & x & x & x & x \\
x & x & x & x & x & x & x & x & x & x & x & x & x & x \\
0 & 0 & 1 & 0 & 0 & x & 0 & 0 & 0 & x & x & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & x & 0 & 0 & 0 & x & x & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & x & 0 & 0 & 0 & x & x & 0 & 1 & 0
\end{bmatrix}
\quad (i\ or\ j = 1\ to\ 14)
$$

Fig. 4 The flow chart of delete repeated nodes and generate new serial nodes number

The serial number of nodes in the initial arbitrary network would be broken, so we must readjustment the sequence of the nodes. We could reuse the delete node number as the next node mark and cycle repeating, then we could get the final incidence matrix and the arbitrary network with new serial number of nodes. The flow chart is shown in Fig. 4. Finally, we can get the adjacent matrix C_A:

$$C_{Aij} = \begin{bmatrix} C_{L1ij} & & & \\ & C_{L2ij} & & \\ & & C_{L3ij} & \\ & & & C_{L4ij} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad (i \ or \ j = 1 \ to \ 11)$$

Fig. 5 Extra examples of arbitrary network

The main algorithms are summarized as follows:

If node $p = $ node q represents that p, q are repeated nodes, then $c'_{pk} = c_{pk}$ **OR** c_{qk}. The "**OR**" operation here is ruled by the Boolean **OR** operation:

when $c_{pk} = 0$ and $c_{qk} = 0$; $c'_{pk} = c_{pk}$ **OR** $c_{qk} = 0$;

when $c_{pk} = 0$ and $c_{qk} = 1$; $c'_{pk} = c_{pk}$ **OR** $c_{qk} = 1$;

when $c_{pk} = 1$ and $c_{qk} = 0$; $c'_{pk} = c_{pk}$ **OR** $c_{qk} = 1$;

when $c_{pk} = 1$ and $c_{qk} = 1$; $c'_{pk} = c_{pk}$ **OR** $c_{qk} = 1$.

where p, q is the serial number of row and column in the matrix; c_{pk} is the initial element in the pth row and the kth column; c'_{pk} is the element after **OR** operations in the pth row and the kth column; k is 1 to n.

Cycle repeating all nodes, we can get the incidence matrix of the arbitrary network. Afterwards, we have to delete the serial number of repeated nodes and readjust the ranks in the matrix.

If node $p = $ node q represents p, q are repeated nodes and $p < q$, then delete the q row and the q column. Meanwhile, the other rows and columns s $(s > q)$ move forward and the nodes serial number renumber; Cycle repeating this, we can get the final incidence matrix C_{Aij}.

4 Examples

The incidence matrix for any arbitrary network can be generated by the processing above. We have been compiled the calculation procedure for it. If we know the line segments' endpoints, the procedure could calculate the intersections between the line segments and generate the incidence matrix. Figure 5a, b are the examples of two arbitrary networks. By the procedure explained above, we could obtain the incidence matrices, which are shown in Figs. 6 and 7. The calculation procedure doesn't move or delete the ranks with the corresponding repeated nodes but instead by "0" in the matrix.

Node	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0	1	0	0	0	0	0	0	0	0	0	0	0
2	1	0	1	0	1	0	1	0	0	0	0	0	0
3	0	1	0	1	0	0	0	1	0	1	0	0	0
4	0	0	1	0	0	0	0	0	0	0	1	0	1
5	0	1	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	1	0	0	0	0	0	0	0	0	0	0	0
8	0	0	1	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	1	0	0	0	0	0	0	0	0	0	0
11	0	0	0	1	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	1	0	0	0	0	0	0	0	0	0

Fig. 6 Calculated results for extra examples as Fig. 5a shown

Node	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	1	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
4	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	1	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	1	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0
9	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	1
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0

Fig. 7 Calculated results for extra examples as Fig. 5b shown

5 Conclusion

A new principle and algorithm for generating the incidence matrix of any arbitrary network were introduced. The calculated procedure was compiled and the calculated results of extra examples verified that the principle and algorithm we stated

were right for auto-generating the incidence matrix for any arbitrary network. It offers fundamental theory and reference tactics to further analyze the path of fluid flow or diffusion in fracture network.

References

 1. Bang-Jensen J, Gutin G (2001) Digraphs. Theory, algorithms and applications, Springer monographs in mathematics. Springer, London
 2. Severini S (2006) On the structure of the adjacency matrix of the line digraph of a regular digraph. Discrete Appl Math 154:1763–1765
 3. Ferrero D (1999) Introduction to interconnection network models, Universitat Politècnica de Catalunya, Espanya. Universidad de la República, Uruguay, pp 1–57
 4. Diestel R (2005), Graph theory, Graduate texts in mathematics, 173, 3rd edn. Springer, London. ISBN 3-540-26183-4
 5. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D-U (2006) Complex networks: structure and dynamics. Phys Rep 424:175–308
 6. Takenaka Y (1970) Graph theoretic concepts and the incidence matrix. Inf Control 17:113–121
 7. Yuan H (2011) Network topology for the application research of electrical control system fault propagation. Procedia Eng 15:1748–1752
 8. Liu G, Sasaki H, Yorino N (2001) Application of network topology to long range composite expansion planning of generation and transmission lines. Electr Power Syst Res 57(3):157–162
 9. Liu L, Zhao X (2012) Application of improved adjacency matrix multiplication in distribution network flow calculation. Energy Procedia 14:983–989
10. Newman MEJ (2003) The structure and function of complex networks. SIAM Rev 45:167–256
11. Estrada E, Rodríguez-Velázquez JA (2006) Subgraph centrality and clustering in complex hyper-networks. Physica A: Stat Mech Appl 364:581–594
12. Mamakani K, Myrvold W, Ruskey F (2012) Generating simple convex Venn diagrams. J Discrete Algorithms 16:270–286
13. van Douwen EK (1993) Applications of maximal topologies. Topology Appl 51(2):125–139
14. Boltyanski V (2005) Application of topology in optimization theory. Topology Applications 146–147:617–628
15. Hoffman AJ, Schieber B (2001) The edge versus path incidence matrix of series-parallel graphs and greedy packing. Discrete Appl Math 113:275–284
16. Chung FRK (1995) Eigenvalues of graphs. In: Proceedings of the international congress of mathematics, 1994. Birkhauser Verlag, New York, pp 1333–1342
17. Griffith DA (2004) Extreme eigenfunctions of adjacency matrices for planar graphs employed in spatial analyses. Linear Algebra Appl 388:201–219
18. Thompson JG (1997) Incidence matrices of finite projective planes and their eigenvalues. J Algebra 191(1):265–278
19. Ilmonen P, Haukkanen P, Merikoski JK (2008) On eigenvalues of meet and join matrices associated with incidence functions. Linear Algebra Appl 429(4):859–874
20. Haemers WH, Omidi GR (2011) Universal adjacency matrices with two eigenvalues. Linear Algebra Appl 435(10):2520–2529

21. Ying K, Chu J, Qu J, Luo Y (2012) A model and topological analysis procedures for a pipeline network of variable connectivity. Adv Eng Softw 48:40–51
22. Ghosh S, Podder M, Sen MK (2010) Adjacency matrices of probe interval graphs. Discrete Appl Math 158:2004–2013

An Efficient and Improved Particle Swarm Optimization Algorithm for Swarm Robots System

Zhiguo Shi, Xiaomeng Zhang, Jun Tu and Zhiyong Yang

Abstract In recent years, the number of researches in which swarm intelligence shown by individual communication in swarm robots is increasing. As one of the representative algorithms in swarm intelligence, particle swarm optimization has been applied to many fields because of its simple concept, easy realizing and good optimization characteristics. However, it still has some disadvantages such as easy falling in the local best situation and solving the discrete optimization problems poor. In this paper, genetic algorithm has been integrated with particle swarm optimization to improve the performance of the algorithm; the simple particle swarm optimization algorithm has been simulated in the Player/Stage and compared with the particle swarm optimization. The simulation shows that the algorithm is faster and more efficient.

Keywords Swarm robots · Particle swarm optimization · Simulation analysis

1 Introduction

Swarm intelligence is a method to achieve artificial intelligence by imitating biological group behavior in the natural world [1], which offers a new thought to the solutions of complex issues by using group advantage without centralized control and global model [2].

Z. Shi (✉) · X. Zhang · J. Tu · Z. Yang
School of Computer and Communication Engineering, University of Science
and Technology Beijing, Beijing, People's Republic of China
e-mail: szg@ustb.edu.cn

X. Zhang
e-mail: zhangxm0703@163.com

J. Tu
e-mail: tujun2011@163.com

The researches of swarm robot systems drew lessons from optimization techniques and principles such as swarm intelligence. It is an application of swarm intelligent in the multi-robot system [3], which is also a new research field in which multi-robot system is given the nature of group in the usual sense [4]. As a representative of the swarm intelligent algorithms, particle swarm algorithm and swarm robots searching are the instances of the smart agent searching. And there is a certain mapping relation between them. Therefore, swarm robots in the real world can be modeled and simulated by the use of particle swarm optimization [5].

In order to improve the convergent speed and accuracy of the algorithm, Genetic Algorithm (GA) has been integrated with PSO in this paper. The simple particle swarm optimization (sPSO) Algorithm which proposed by Hu Wang et al. combined with obstacle avoidance principle has been simulated in the Player/Stage and compared with the bPSO in the timeline.

2 Research Background and Related Work

Genetic algorithm, one of the three major branches of evolutionary computation, is a kind of adaptive, probabilistic, random and iterative search algorithm. It is evolved through the reference from the evolution law of the biosphere. Particle swarm algorithm is used to solve continuous optimization problems originally. Particles in the solution space search the goal through following the best particle. Therefore, it achieves simply and has fewer parameters to be adjusted.

2.1 Genetic Algorithm

Genetic algorithm is a new global optimization algorithm which develops in recent years [6]. Professor Holland first put forward GA in 1962, the idea of which is based on Darwin's evolution and Mendel's genetic laws. The most important content of Darwin's evolution is the survival of the fittest, which observes that the population in each generation always developing towards advance and increasingly adapting to the environment. The most important content of Mendel's genetic laws is the theory of genetics, which argues that Genetic exists as codes in cells and as genes contained inside the chromosome.

The process of GA is in the following.

(a) Encode and generate the initial population

Select the appropriate encoding method, and randomly generate initial population with certain length formed by N chromosomes.

$$pop_i(t), t = 1, i = 1, 2, 3, \ldots, N.$$

(b) Calculate the fitness

Calculate the fitness of each chromosome *popi(t)* in the population *pop(t)*.

$$f_i = fitness(pop_i(t)).$$

(c) Determine whether the convergence criteria of the algorithm is met. If it is, output search results. Otherwise, continue to perform the following steps.

(d) Selection operation. Calculate the probability of selection in accordance with the fitness of each individual.

$$p_i = \frac{f_i}{\sum\limits_{i=1}^{N} f_i}, \quad i = 1, 2, 3, \ldots, N$$

(e) Randomly select some chromosomes from the population of the current generation which to form a new population in the population of the next generation in accordance with the above probability distribution.

$$newpop(t+1) = \{pop_i(t)|j = 1, 2, \ldots, N\}$$

(f) Crossover operation. With the crossover probability *Pc*, cross over the pair at a point chosen randomly to form two new strings. Then A population *cross-pop(t + 1)* composed of *N* chromosomes will be obtained.

(g) Mutation operation. With a smaller probability *Pm*, make the genes of chromosomes mutate to form a new population *mutpop(t + 1)*. It is not only the offspring recorded as *pop(t) = mutpop(t + 1)* obtained from finishing the mutation operation, but also the parent of the next genetic manipulation. Return (b).

2.2 Particle Swarm Optimization

The particle swarm optimization (PSO) is first put forward by Kennedy and Eberhart in [7, 8], which is a kind of heuristic search algorithm based on population optimization. The basic idea is originated from the research of bird flock foraging and shoal learning behavior.

The principle of the particle swarm algorithm is: there is only a piece of food in the region, and all the birds don't know where the food is, but they know the distance between the current location and the food. What is the optimal strategy to find the food? The most simple and effective method is to search the area around the bird whose location is nearest from the food at present.

In the PSO, each solution of optimization problem corresponding to the location of a bird in the search space, calling the bird "Particle". D-dimensional position vector of the ith particle is $x_i = (x_{i1}, x_{i2}, \ldots, x_{iD})$. According to pre-set fitness functions (which is relevant to the problem), the current value of xi can be

calculated. Particles fly at a certain speed in the search space. $v_i = (v_{i1}, v_{i2}, \ldots, v_{id}, \ldots, v_{iD})$ is the speed of particle i, which dynamically adjust based on its own flying experience and that of companion. All the particles have a Fitness Value decided by target function, and know the best position so far (particle best, notes for pbest) $P_i = (P_{i1}, P_{i2}, \ldots, P_{id}, \ldots, P_{iD})$ and the current position. These can be seen as their own flying experience. In addition, each particle also know the best position (global best, notes for gbest) so far in the group, which can be regarded as the companion's experience of a particle. And then the particles follow the current optimal particle to search t in the solution space.

The velocity and position of the particle in each iteration are updated based on the following formula:

$$v_{id}^{k+1} = \omega^* v_{id}^k + c_1 r_1 \left(p_{id} - x_{id}^k \right) + c_2 r_2 \left(p_{gd} - x_{id}^k \right) \tag{1}$$

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} \tag{2}$$

k is the number of iterations, w is the inertia weight, r_1 and r_2 are the random numbers between [0, 1], which are used to keep the diversity of the population. c_1 and c_2 are learning factors, which make particles have the ability to summarize themselves and learn from excellent individual in the group, thus close to pbest and gbest.

2.3 A Simpler and More Effective Particle Swarm Optimization Algorithm

PSO and most of the improved algorithms all based on the two key factors of particles, "position" and "speed". Therefore, the equations of the improved algorithms all contain a position variable and the speed variable. For most of the improved algorithms of PSO, Some operators like crossover, mutation, etc. are added, which makes the description of PSO more and more complicated. This also makes the quantitative analysis of the convergence of PSO very complicated.

We can found something from the analysis of the biological model of PSO and its evolutionary iteration equation. In PSO, the velocity variable of the particle is not required. Viewed from the point of the model of the basic particle swarm optimization (bPSO), the position xi represents the solution of this problem. Because the final result of the optimization is xi infinitesimal approaching to the optimal solution, it is only need only consider the direct change of x_i. The Velocity vector v_i just represents the rapidity of the particles' movement. The speed of the movement is not able to show that the particle can approaches to the location of the optimal solution effectively. Instead, that may cause the particles to deviate from the correct direction of evolution, which is the "divergence" phenomenon, thus resulting in the phenomenon of slow convergence speed and low convergence precision in the later stage. In addition, the position and velocity directly are

computed in additions without the concept of particles' movement time. It is not in line with the law of motion x = vt in reality life.

Hu and Li [9] proposed an optimization equation of the particle swarm without speed items.

$$x_{id}^{k+1} = wx_{id}^k + c_1 r_1 \left(p_{id} - x_{id}^k \right) + c_2 r_2 \left(p_{gd} - x_{id}^k \right) \tag{3}$$

wx_{id}^k in the formula represents the position of the particles in the previous stage, called the "Historical" part. It carries the impact of the past. Through changing the value of w, we control the degree of the impact of the next position. $c_1 r_1 (p_{id} - x_{id}^k)$ is the difference value between the historical best position of the particle and the current position of the particle. It makes the particle have the trend to the historical best position, called the "Acknowledge" part. It expresses the thoughts of the particle on itself. $c_2 r_2 (p_{gb} - x_{id}^k)$ is the difference value between the historical best position of the group and the current position of the group. It makes the group have the trend to the historical best position, called the "Social Experience" part. It expresses the comparison and the imitation between the particles of the group, and achieves the information sharing and collaboration between particles. This is the simple particle swarm optimization, abbreviated as sPSO.

3 An Efficient and Improved Particle Swarm Optimization Algorithm

The main defects of particle swarm algorithm are its local searching capability which is poor and its precocious phenomena. The reduced diversity of the population in the search space causes PSO to trap into a local optimum. The solutions diversity is improved owing to the different evolutionary patterns, crossover and mutation, of GA. But GA often leads to inactive redundancy iteration, long computation time and reduced efficiency of solving reduce when it runs to a certain range. Therefore, the subject improves the search performance of PSO by combining PSO and GA. It is the genetic and basic particle swarm optimization (gbPSO).

There are some particles are closer to the optimal solution than the other particles after being initialized in general. It is easy to think that the probability of finding the optimization solution will be biggish if these particles mutate. The idea of improved PSO algorithm is:

1. Divide a population which has n individuals into two groups, good group and poor group, in certain proportion (assuming 1:1 or depending on the issue and population size). If the termination condition is achieved, stop running the algorithm. Otherwise, continue with the steps below. Evaluate each individual and arrange them in quality. The better among them constitute the good group, the number individuals of which is n/2. The rest constitute the poor group.
2. Find gbest in the whole group and eliminate the poor group.
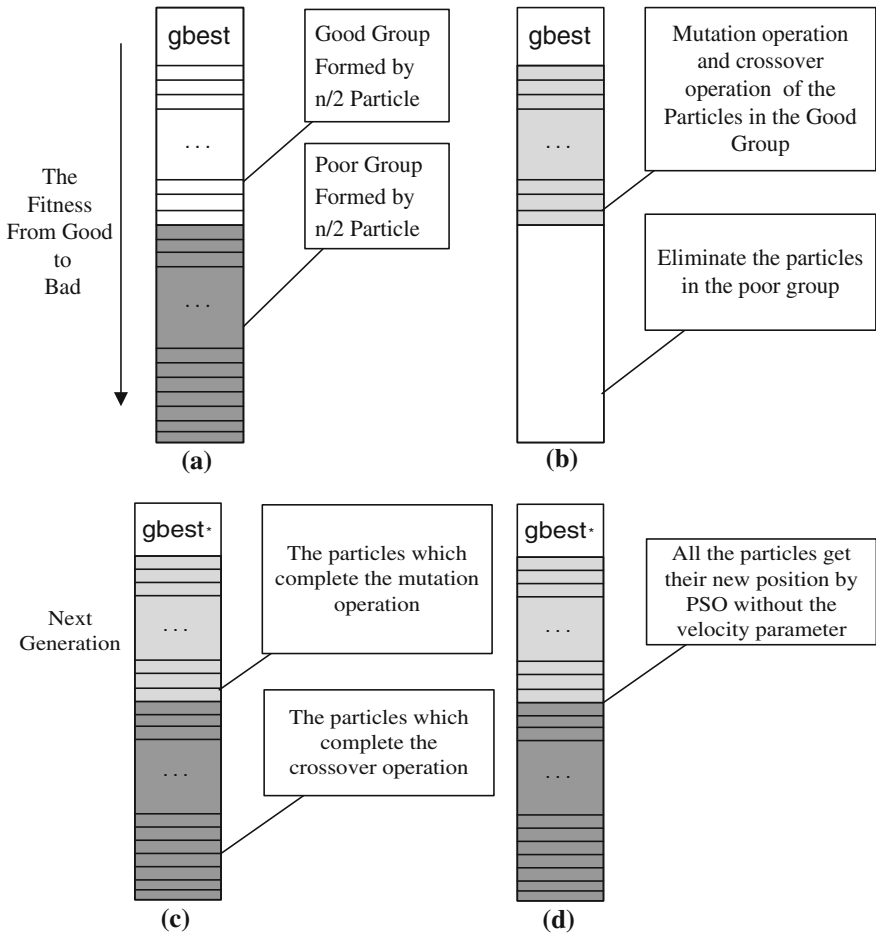
Fig. 1 a–d The idea of the improved algorithm

3. The particles in the good group do the mutation operation and crossover operation respectively. Compare these particles with gbest. If better, update gbest.
4. The position of all the particles are updated based on the following formula which is similar to (3). Then, return (1).

$$x_{id}^{k+1} = w x_{id}^{k} {}^{*} rand \; + c_1 r_1 \left(p_{id} - x_{id}^{k} \right) + c_2 r_2 \left(p_{gd} - x_{id}^{k} \right) \tag{4}$$

What shown in Fig. 2 is helpful in understand the idea of gaPSO (Fig. 1).

4 Simulation and Analysis

4.1 Experiment Design and Results Analysis

In order to test the performance of gaPSO proposed in this paper, two benchmark functions which can be found in [9] are used in the experiment. They are commonly used in the comparison of the optimization algorithms. Function form, dimension, search range and the extremum in theory and are shown in Table 1.

The number of particles is 16, the number of iterations is 150 and final results is the average value of the results after independently running 20 times. $c_1 = c_2 = 2$, $\omega = 0.4$–0.9.

The simulation results are shown in Table 2.

From Table 2 we can see that gaPSO has better convergence precision than bPSO. Although the average value of its simulation results for $f2$ is 2.368675E-15, it finds the optimal value 0 seventeen times in the twenty simulations. From the perspective of the length of simulation time, gaPSO may not be the best choice, but its result is in the acceptable range. The advantage of gaPSO is obvious.

4.2 Simulation in Player/Stage for Swarm Robots System

Player/Stage simulation platform is an open experimental platform, which has been used, modified and extended by many researchers in the robots field. It can provide enough simulation environments, and if the simulation capability of the stage is strong enough, the client terminal will not be able to distinguish the real robot equipment and the Stage alternatives—Virtual Device, achieves the testing efforts.

The simulation scenario in this article is designed as a common university campus. The search target is the larger dots. The number of the robots, the small dots, is 20. The task of the swarm robots is to successful avoiding the obstacles and finding the target dot.

The following is a series of simulation diagram.
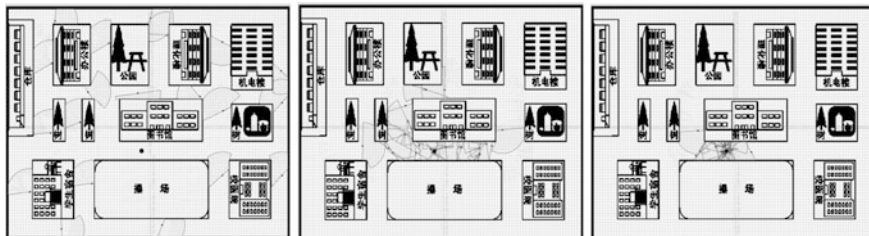
Fig. 2 Simulation diagram

Table 1 Benchmark functions used to test the improvement methods

Name and code	Formula	Dim n	Range $[x_{min}, x_{max}]$	Optimal f
Sphere $f1$	$f_1(x) = \sum_{i=1}^{n} x_i^2$	30	$[-100, 100]^{30}$	0
Rastrigin $f2$	$f_2(x) = \sum_{i=1}^{n} \left(x_i^2 - 10\cos(2\pi x_i) + 10 \right)$	30	$[-100, 100]^{30}$	0

Table 2 Simulation results

	Algorithm	
	bPSO	gaPSO
$f1$ time (s)	0.028846	0.773109
$f1$ average Value	2.239555	1.24E-78
$f2$ time (s)	0.037629	1.107266
$f2$ average value	28.787979	2.368675E-15

The size of the environment is set in the file called "world" in Player/Stage. The name, color, initial position and posture of the 20 pioneer robots with laser sensors are set in configuration.

During the pre-run of the improved algorithm, s since no particles find the target, the convergence speed is relatively slow. During the mid-part of the improved algorithm, since some robots have found the target, the others converge more quickly.

In the last runs of the algorithm, most of the particles have found the target, and the others quickly converge to the position of the target.

5 Conclusion

Particle swarm optimization is an evolution computing technology based on swarm intelligence method. In this paper, the idea of an algorithm called genetic and basic particle swarm optimization is proposed to make up the disadvantages. The simulations of bPSO and gaPSO show the effectiveness and feasibility of gaPSO. It can be found from the analysis that gaPSO improved the precision of convergence, but the speed is significantly lower than the bPSO.

References

1. Kennedy J, Eberhart RC (2001) Swarm intelligence. Academic Press, USA
2. Wang M, Zhu YL, He XX (2005) Research summarize of swarm intelligence. Comput Proj 31:194–196
3. Xue SD, Zeng JC (2008) Swarm robotics: a survey. Pattern Recognit Artif Intell l21:177–185
4. Bayindir L, Sahin E (2007) A review of studies in swarm robotics. Turk J Electr Eng Comput Sci 15:115–147
5. Chen BD (2009) Improved particle swarm algorithm based on the characteristics of swarm robots. Taiyuan University of Science and Technology, Taiyuan
6. Liu Y (2003) Non-numerical parallel algorithms (volume II) genetic algorithm. Science Publishing House, Beijing, pp 1–6
7. Kennedy J, Eberhart R (1995) Particle swarm optimization. In: Proceedings of IEEE international conference on neural networks, Washington, pp 1942–1948
8. Eberhart R, Kennedy J (1995) A new optimizer using particle swarm theory. In: Proceedings of the 6th international symposium on micro machine and human science, Nagoya, pp 39–43
9. Hu W, Li ZS (2007) A simpler and more effective particle swarm optimization algorithm. J Softw 18:861–868

Ratio Estimation and Regression Estimation Method in Three-Stage Sampling

Shu Lv and Bing Chen

Abstract This paper discusses the ratio estimator and regression estimator of population mean in the three-stage sampling, in which all level of unit sizes are equal, when some auxiliary information can be available.

Keywords Three-stage sampling · Auxiliary information · Ratio estimator · Regression estimator

1 Introduction

Multi-stage sampling is widely applied due to its flexibility and saving manpower, material, financial resources and time. There have been abundant theories about the single-stage and two-stage sampling [1–3]. However, in practical work, the population contains a lot of units, and distributes widely. It is difficult for the single-stage and two-stage sampling to produce an overall sample, which needs to take samples by three-stage sampling. Ye [4] has given the simple valuation method in three-stage sampling. This paper proposes the ratio estimator and regression estimator of population mean in three-stage sampling, in the situation that the related auxiliary information is available.

S. Lv (✉) · B. Chen
School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, People's Republic of China
e-mail: lvshu@uestc.edu.cn

B. Chen
e-mail: chenbingbingjiayou@yahoo.cn

2 Some Notations About Estimation Problem

Dispose the population contains N primary units, each primary unit contains M secondary units, and each secondary unit contains K third-level units. Sampling is divided into three stages by using simple random sampling method, and the sample size obtained are respectively n, m and k. And in the sampling survey the information of auxiliary variable X, which has the positive correlation relationship with Y, can be obtained. Thus, introduce the following notations.

Let Y_{iju}, X_{iju} denote the survey index and auxiliary index value of the uth third-level unit of the jth secondary unit of the ith primary unit in the population unit; While y_{iju}, x_{iju} denote the sample of survey index, auxiliary index value. $\overline{Y}_{ij} = \frac{1}{K} \sum_{u=1}^{K} Y_{iju}$, $\overline{X}_{ij} = \frac{1}{K} \sum_{u=1}^{K} X_{iju}$ represent the survey index and auxiliary index mean of the population secondary unit to the third-level unit; $\overline{y}_{ij} = \frac{1}{k} \sum_{u=1}^{k} y_{iju}$, $\overline{x}_{ij} = \frac{1}{k} \sum_{u=1}^{k} x_{iju}$ represent the survey index and auxiliary index mean of the sample secondary unit to the third-level unit; $\overline{\overline{Y}}_i = \frac{1}{M} \sum_{j=1}^{M} \overline{Y}_{ij}$, $\overline{\overline{X}}_i = \frac{1}{M} \sum_{j=1}^{M} \overline{X}_{ij}$ denote the survey index and auxiliary index mean of the population primary unit to the third-level unit; $\overline{\overline{y}}_i = \frac{1}{m} \sum_{j=1}^{m} \overline{y}_{ij}$, $\overline{\overline{x}}_i = \frac{1}{m} \sum_{j=1}^{m} \overline{x}_{ij}$ denote the survey index and auxiliary index mean of the sample primary unit to the third-level unit; $\overline{\overline{\overline{Y}}} = \frac{1}{N} \sum_{i=1}^{N} \overline{\overline{Y}}_i$, $\overline{\overline{\overline{X}}} = \frac{1}{N} \sum_{i=1}^{N} \overline{\overline{X}}_i$ signify the survey index and the auxiliary index average of the third-level unit of the population, and $\overline{\overline{\overline{y}}} = \frac{1}{n} \sum_{i=1}^{n} \overline{\overline{y}}_i = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \overline{y}_{ij}$, $\overline{\overline{\overline{x}}} = \frac{1}{n} \sum_{i=1}^{n} \overline{\overline{x}}_i = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \overline{x}_{ij}$ denote the survey index and auxiliary index average of the sample third-level unit respectively. The sampling ratios of each stage are $f_1 = \frac{n}{N}$, $f_2 = \frac{m}{M}$, $f_3 = \frac{k}{K}$. The variance among the population primary units is S_1^2, and among the secondary units is S_2^2. $R_{ij} = \frac{Y_{ij}}{X_{ij}} = \frac{\overline{Y}_{ij}}{\overline{X}_{ij}}$ shows the ratio of the jth sub-unit of the ith primary unit of the population. Finally, $R = \frac{Y}{X} = \frac{\overline{\overline{\overline{Y}}}}{\overline{\overline{\overline{X}}}}$ represents the population ratio.

3 Ratio Estimation Method

Two kinds of ratio estimators are given as followings.

3.1 Separate Ratio Estimator

Set $\overline{y}_{Rij} = \frac{\overline{y}_{ij}}{\overline{x}_{ij}} \overline{X}_{ij}$, then $\overline{y}_{Rij}$ is the ratio estimator of $\overline{Y}_{ij}$. So we estimate population mean $\overline{\overline{\overline{Y}}}$ by

$$\bar{\bar{\bar{y}}}_{RS} = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \bar{y}_{Rij}. \tag{1}$$

$\bar{\bar{\bar{y}}}_{RS}$ is known as the separate ratio estimator of the population mean $\bar{\bar{\bar{Y}}}$.

According to the nature of the ratio estimator [5], when the sample size k is large enough, there is the expectation of $\bar{\bar{\bar{y}}}_{RS}$. $E\left(\bar{\bar{\bar{y}}}_{RS}\right) \approx \bar{\bar{\bar{Y}}}$, So $\bar{\bar{\bar{y}}}_{RS}$ is an approximately unbiased estimator of the population mean, when the sample size is large enough.

According to the variance formula of the three-step sampling estimator [6] and the nature of the ratio estimator, the variance of $\bar{\bar{\bar{y}}}_{RS}$ is

$$V\left(\bar{\bar{\bar{y}}}_{RS}\right) = \frac{1-f_1}{n} S_1^2 + \frac{1-f_2}{nm} S_2^2 + \frac{1-f_3}{nmk} \frac{\sum_{i=1}^{N} \sum_{j=1}^{M} \sum_{u=1}^{K} \left(Y_{iju} - R_{ij} X_{iju}\right)^2}{NM(K-1)}.$$

When the survey index Y_{iju} and auxiliary index X_{iju} is roughly in direct proportion, that is $Y_{iju} - R_{ij} X_{iju} \approx 0$, the third formula on the right side of the above equation can be omitted. And then the following formula will set up.

$$V\left(\bar{\bar{\bar{y}}}_{RS}\right) \approx \frac{1-f_1}{n} S_1^2 + \frac{1-f_2}{nm} S_2^2$$

3.2 Combined Ratio Estimator

Define the combined ratio estimator of the population mean $\bar{\bar{\bar{Y}}}$ is

$$\bar{\bar{\bar{y}}}_{RC} = \frac{\bar{\bar{\bar{y}}}}{\bar{\bar{\bar{x}}}} \bar{\bar{\bar{X}}} = \frac{y}{x} \bar{\bar{\bar{X}}}.$$

In order to study the unbiasedness of $\bar{\bar{\bar{y}}}_{RC}$, the following should be taken into consideration

$$\bar{\bar{\bar{y}}}_{RC} - \bar{\bar{\bar{Y}}} = \frac{\bar{\bar{\bar{y}}}}{\bar{\bar{\bar{x}}}} \bar{\bar{\bar{X}}} - R\bar{\bar{\bar{X}}} = \frac{\bar{\bar{\bar{X}}}}{\bar{\bar{\bar{x}}}} \left(\bar{\bar{\bar{y}}} - R\bar{\bar{\bar{x}}}\right). \tag{2}$$

Because of $E\left(\bar{\bar{\bar{x}}}\right) = \bar{\bar{\bar{X}}}$, and when the sample size k is large enough, there is $\bar{\bar{\bar{x}}} \approx \bar{\bar{\bar{X}}}$, thus $\bar{\bar{\bar{y}}}_{RC} - \bar{\bar{\bar{Y}}} \approx \bar{\bar{\bar{y}}} - R\bar{\bar{\bar{x}}}$. Then $E\left(\bar{\bar{\bar{y}}}_{RC}\right) \approx \bar{\bar{\bar{Y}}}$. So, when the sample size is large enough, $\bar{\bar{\bar{y}}}_{RC}$ is an approximately unbiased estimator of the population mean $\bar{\bar{\bar{Y}}}$.

The variance of $\bar{\bar{\bar{y}}}_{RC}$ is as follows

$$V\left(\bar{\bar{\bar{y}}}_{RC}\right) = E\left[\bar{\bar{\bar{y}}}_{RC} - E\left(\bar{\bar{\bar{y}}}_{RC}\right)\right]^2 \approx E\left[\bar{\bar{\bar{y}}}_{RC} - \bar{\bar{\bar{Y}}}\right]^2 \approx E\left(\bar{\bar{\bar{y}}} - R\bar{\bar{\bar{x}}}\right)^2. \tag{3}$$

Let $G_{iju} = Y_{iju} - RX_{iju}, \quad 1 \le i \le N, 1 \le j \le M, 1 \le u \le K.$

Then $\overline{G}_{ij} = \overline{Y}_{ij} - R\overline{X}_{ij} \quad \overline{\overline{G}}_i = \overline{\overline{Y}}_i - R\overline{\overline{X}}_i \quad \overline{\overline{\overline{G}}} = \overline{\overline{\overline{Y}}} - R\overline{\overline{\overline{X}}} = 0.$

And let $g_{iju} = y_{iju} - Rx_{iju} \quad , \overline{g}_{ij} = \overline{y}_{ij} - R\overline{x}_{ij} \quad \overline{\overline{g}}_i = \overline{\overline{y}}_i - R\overline{\overline{x}}_i \quad \overline{\overline{\overline{g}}} = \overline{\overline{\overline{y}}} - R\overline{\overline{\overline{x}}}.$

Then $E_3\left(\overline{g}_{ij}\right) = \overline{G}_{ij}, E\left(\overline{\overline{\overline{g}}}\right) = \overline{\overline{\overline{G}}} = 0.$

Thus

$$V\left(\overline{\overline{\overline{y}}}_{RC}\right) \approx \frac{1-f_1}{n} \frac{\sum_{i=1}^{N}\left(\overline{\overline{Y}}_i - R\overline{\overline{X}}_i\right)^2}{N-1} + \frac{1-f_2}{nm} \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\left[\left(\overline{Y}_{ij} - \overline{\overline{Y}}_i\right) - R\left(\overline{X}_{ij} - \overline{\overline{X}}_i\right)\right]^2}{N(M-1)}$$

$$+ \frac{1-f_3}{nmk} \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{u=1}^{K}\left[\left(Y_{iju} - \overline{Y}_{ij}\right) - R\left(X_{iju} - \overline{X}_{ij}\right)\right]^2}{NM(K-1)}$$

When the three-stage sample size k is large enough, the last term on the right side of the above equation can be omitted. Now

$$V\left(\overline{\overline{\overline{y}}}_{RC}\right) \approx \frac{1-f_1}{n} \frac{\sum_{i=1}^{N}\left(\overline{\overline{Y}}_i - R\overline{\overline{X}}_i\right)^2}{N-1}$$

$$+ \frac{1-f_2}{nm} \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\left[\left(\overline{Y}_{ij} - \overline{\overline{Y}}_i\right) - R\left(\overline{X}_{ij} - \overline{\overline{X}}_i\right)\right]^2}{N(M-1)}. \qquad (4)$$

4 Regression Estimation Method

Two kinds of regression estimation of the three-stage sampling will be displayed as followings.

4.1 Separate Regression Estimator

Define $\overline{y}_{lrij} = \overline{y}_{ij} + \beta_{ij}\left(\overline{X}_{ij} - \overline{x}_{ij}\right)$, $\overline{y}_{lrij}$ as the regression estimator of $\overline{y}_{ij}$, and β_{ij} denotes regression coefficient. So the separate regression estimator of population mean $\overline{\overline{\overline{Y}}}$ is

$$\overline{\overline{\overline{y}}}_{lrs} = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\overline{y}_{lrij}. \qquad (5)$$

When the regression coefficient β_{ij} is given in advance as β_0.

There is $E_3\left(\overline{y}_{lrij}\right) = \overline{Y}_{ij}$, and thus the expectation of $\overline{\overline{\overline{y}}}_{lrs}$ is $E\left(\overline{\overline{\overline{y}}}_{lrs}\right) = \overline{\overline{\overline{Y}}}$. So $\overline{\overline{\overline{y}}}_{lrs}$ is the unbiased estimator of the population mean $\overline{\overline{\overline{Y}}}$.

Then, the variance of $\overline{\overline{\overline{y}}}_{lrs}$ is

$$V\left(\overline{\overline{\overline{y}}}_{lrs}\right) = \frac{1-f_1}{n}S_1^2 + \frac{1-f_2}{nm}S_2^2$$
$$+ \frac{1-f_3}{nmk}\frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{u=1}^{K}\left[\left(Y_{iju} - \overline{Y}_{ij}\right) - \beta_0\left(X_{iju} - \overline{X}_{ij}\right)\right]^2}{NM(K-1)}.$$

When the survey index Y_{iju} and auxiliary index X_{iju} has linear regression relation, the third term on the right side of the above equation can be omitted. Then

$$V\left(\overline{\overline{\overline{y}}}_{lrs}\right) \approx \frac{1-f_1}{n}S_1^2 + \frac{1-f_2}{nm}S_2^2.$$

When the regression coefficient β_{ij} cannot be set in advance.

When the regression coefficient β_{ij} cannot be set in advance, we need to estimate β_{ij} by sample regression coefficient b_{ij}, where

$$b_{ij} = \frac{\sum_{u=1}^{k}\left(y_{iju} - \overline{y}_{ij}\right)\left(x_{iju} - \overline{x}_{ij}\right)}{\sum_{u=1}^{k}\left(x_{iju} - \overline{x}_{ij}\right)^2}.$$

At this time, when the linear regression relationship is established, and k is large enough

$$E\left(\overline{\overline{\overline{y}}}_{lrs}\right) \approx \overline{\overline{\overline{Y}}}, \quad V\left(\overline{\overline{\overline{y}}}_{lrs}\right) \approx \frac{1-f_1}{n}S_1^2 + \frac{1-f_2}{nm}S_2^2.$$

4.2 Combined Regression Estimator

Define

$$\overline{\overline{\overline{y}}}_{lrc} = \overline{\overline{\overline{y}}} + \beta\left(\overline{\overline{\overline{X}}} - \overline{\overline{\overline{x}}}\right) = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\left[\overline{y}_{ij} + \beta\left(\overline{\overline{\overline{X}}} - \overline{x}_{ij}\right)\right].$$

Then $\overline{\overline{\overline{y}}}_{lrc}$ is the combined regression estimator of population mean $\overline{\overline{\overline{Y}}}$.

When β is set as constant in advance

$$E\left(\overline{\overline{\overline{y}}}_{lrc}\right) = E\left[\overline{\overline{\overline{y}}} + \beta\left(\overline{\overline{\overline{X}}} - \overline{\overline{\overline{x}}}\right)\right] = E\left(\overline{\overline{\overline{y}}}\right) + \beta\left[\overline{\overline{\overline{X}}} - E\left(\overline{\overline{\overline{x}}}\right)\right] = \overline{\overline{\overline{Y}}}. \tag{6}$$

Thus $\overline{\overline{\overline{Y}}}$ is the unbiased estimator of the population mean $\overline{\overline{\overline{Y}}}$.

According to the variance formula of the three-step sampling and the nature of the regression estimator, the variance of $\overline{\overline{\overline{y}}}_{lrc}$ is

$$V\left(\overline{\overline{\overline{y}}}_{lrc}\right) = \frac{1-f_1}{n} \frac{\sum_{i=1}^{N}\left[\left(\overline{\overline{Y}}_i - \overline{\overline{\overline{Y}}}\right)+\beta\left(\overline{\overline{X}} - \overline{\overline{X}}_i\right)\right]^2}{N-1} + \frac{1-f_2}{nm} \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\left[\left(\overline{Y}_{ij} - \overline{\overline{Y}}_i\right)+\beta\left(\overline{X}_{ij} - \overline{\overline{X}}_i\right)\right]^2}{N(M-1)}$$

$$+ \frac{1-f_3}{nmk} \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{u=1}^{K}\left[\left(Y_{iju} - \overline{Y}_{ij}\right)-\beta\left(X_{iju} - \overline{X}_{ij}\right)\right]^2}{NM(K-1)}.$$

$$(7)$$

When the regression coefficient β cannot be set in advance.

When the regression coefficient β cannot be set in advance, we need to estimate β by sample regression coefficient b, where

$$b = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{u=1}^{k}\left(y_{iju} - \overline{y}_{ij}\right)\left(x_{iju} - \overline{x}_{ij}\right)}{\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{u=1}^{k}\left(x_{iju} - \overline{x}_{ij}\right)^2}. \tag{8}$$

References

1. Cochran WG (1977) Sampling techniques. Wiley, New York
2. Yu C (2006) Ratio estimation and regression estimation method in two-stage sampling. Stat Decis 1:24–27
3. Yu C (2007) The estimator of population ratio in view of the two-stage sampling. J Appl Probab Stat 3:319–328
4. Ye A (1996) Simple valuation method in three or four-stage sampling. J Fuzhou Univ 4:35–38
5. Jin Y, Du Z, Jiang Y (2008) Sampling technique. China Renmin University Press, Beijing
6. Du Z (2005) Sampling techniques and practices. Tsinghua University Press and Springer, Beijing

Oscillation Criteria for Second Order Functional Differential Equation

Nan Tang and Jie Zhang

Abstract In this paper, by introducing nonnegative kernel function H(t, s) and h(t, s), using the generalized Riccati technique and the integral averaging technique, second order functional differential equations with deviating arguments are discussed.

1 Introduction

The study of the oscillatory behavior of the solutions of differential equations [1, 2], besides its theoretical interest, is important from the viewpoint of applications. For example, differential equations arise frequently in many applications such as population growth models, control problem. In recent years, the number of investigations devoted to the oscillation theory of the differential equations has increased considerably [3–5].

N. Tang (✉)
College of Science, Anhui University of Science and Technology, Huainan 232001, China
e-mail: tnya369@163.com

J. Zhang
Huaihu Coal and Electric Power Company, Huainan 232001, China
e-mail: ne.kt@163.com

2 Main Results

In this section, we establish oscillation criteria for second order functional differential equation with continuous deviating arguments.

$$\left[r(t)|x'(t)|^{\alpha-1}x'(t)\right]' + \int_a^b q(t,\xi)|x[g(t,\xi)]|^{\alpha}\mathrm{sgn}\, x d\sigma(\xi) = 0, t \geq t_0 \qquad (1)$$

where $\alpha > 0$ is a constant; $r(t) \in C'([t_0,\infty);(0,\infty))$ is nondecreasing in t; $\varphi_a(u) = |u|^{\alpha-1}u; q(t,\xi) \in C([t_0,\infty) \times [a,b], R_+), R_+ = [0,\infty), q(t,\xi) \neq 0, g(t,\xi) \in C([t_0,\infty) \times [a,b], R), g(t,\xi) \leq t, \xi \in [a,b], g(t,\xi)$ is nondecreasing with respect to t, ξ, respectively; $\lim\min_{t\to\infty,\xi\in[a,b]}\{g(t,\xi)\} = \infty; \sigma(\xi) \in C([a,b], R), \sigma(\xi)$ is nondecreasing in ξ, and integral in Eq. (1) is a Stieltjes integral.

Definition 1 A nontrivial solution of Eq. (1) is called oscillatory if it has arbitrarily large zeros, otherwise it is called non-oscillatory.

Definition 2 If all its solutions are oscillatory, Eq. (1) is oscillatory. In order to prove our theorems we use the following well-known inequality.

Lemma 1 [5] *If A, B are nonnegative, then*

$$A^\lambda + (\lambda - 1)B^\lambda - \lambda AB^{\lambda-1} \geq 0, \quad \lambda > 1,$$

and the equality holds if and only if $A = B$.

By introducing nonnegative kernel function H(t, s) and h(t, s), we establish oscillation criteria for Eq. (1).

Theorem 1 *Assume $\frac{d}{dt}g(t,a)$ exists. Let $D = \{(t,s)|t \geq s \geq t_0\}, D_0 = \{(t,s)|t > s \geq t_0.\}; H(t,s) \in C(D,R), h(t,s) \in C(D_0, R_+)$, satisfy the following conditions:*

$$H(t,t) = 0, t \geq t_0; \frac{\partial H(t,s)}{\partial s} \leq 0; H(t,s) > 0; (t,s) \in D_0 \qquad (2)$$

$$h(t,s) = -\frac{\partial H(t,s)}{\partial s}, (t,s) \in D_0 \qquad (3)$$

If

$$\limsup_{t\to\infty} \frac{1}{H(t,t_0)} \int_{t_0}^t \left[H(t,s) \int_a^b q(s,\xi)d\sigma(\xi) - \frac{r(s)|h(t,s)|^{a+1}}{(a+1)^{a+1}[H(t,s)g'(s,a)]^a}\right] ds$$
$$= \infty$$

$$(4)$$

then Eq. ([1](1)) *is oscillatory.*

Proof Let $x(t)$ be a non-oscillatory solution of Eq. ([1](1)). Suppose that $x(t) > 0$, $t \geq t_0$. From $\lim\limits_{t\to\infty, \xi\in[a,b]} \min \{g(t,\xi)\} = \infty$, there exists a $T_0 \geq t_0$, such that $x(t) > 0, x[g(t,\xi)] > 0$, $t \geq T_0, \xi \in [a,b]$. Hence, we have $\text{sgn} x > 0$, $[r(t)\varphi_a(x')]' \leq 0, t \geq T_0$. According to the fact of $r(t)\varphi_a(x')$ is decreasing and $x'(t) > 0$ can be proved. In fact, assume that $x'(t) > 0$ is not true, then there exists a $T \geq T_0$, such that $x'(T) < 0$. Since $(r(t)\varphi_a(x'))' = ar(t)(-x'(t))^{a-1}$ $x''(t) - r'(t)|x'(t)|^a \leq 0$, we get $x''(t) \leq 0$. According to the fact of $x'(t)$ is decreasing, there exists a $t_1 \geq T$, such that $x'(t_1) < 0$, and $x'(t) \leq x'(t_1) < 0, t \geq t_1$. Integrating from t_1 to t, we have $x(t) \leq x(t_1) + x'(t_1) \ (t - t_1)$. Thus, we conclude that $\lim\limits_{t\to\infty} x(t) = -\infty$, this contradicts $x(t) > 0$, $t \geq t_0$. Using $|x(t)|^a$ is nondecreasing and $x'(t)$ is nonincreasing, we get $|x[g(t,a)]|^a \leq |x[g(t,\xi)]|^a$ and $x'(t) \leq x'[g(t,\xi)] \leq x'[g(t,a)]$.

Let

$$\omega(t) = \frac{r(t)\varphi_a(x')}{|x[g(t,a)]|^a}, \tag{5}$$

Then differentiating ([5](5)) and making use of ([1](1)), it follows that

$$\omega'(t) = \frac{r'(t)}{r(t)}\omega(t) + \frac{-\int_a^b q(t,\xi)\varphi_a\{x[g(t,\xi)]\}d\sigma(\xi) - r'(t)\varphi_a(x')}{\varphi_a\{x[g(t,a)]\}}$$
$$- \frac{a\omega(t)x'[g(t,a)]g'(t,a)}{x[g(t,a)]}$$

$$\omega'(t) \leq -\frac{\varphi_a\{x[g(t,a)]\}\int_a^b q(t,\xi)d\sigma(\xi)}{\varphi_a\{x[g(t,a)]\}} - \frac{a\omega(t)x'(t)g'(t,a)}{x[g(t,a)]}$$
$$= -\int_a^b q(t,\xi)d\sigma(\xi) - \frac{a|\omega(t)|^{\frac{a+1}{a}}g'(t,a)}{[r(s)]^{\frac{1}{a}}}, \quad t \geq T.$$

Multiplying, with t replaced by s, by $H(t,s)$, and integrating from T to t, we have $\int_T^t \omega'(s)H(t,s)ds \leq -\int_T^t H(t,s)\int_a^b q(s,\xi)d\sigma(\xi)ds - \int_T^t \frac{aH(t,s)g'(s,a)|\omega(t)|^{\frac{a}{a+1}}}{[r(s)]^{\frac{1}{a}}}ds$

Now integrate the first integral, to obtain

$$\int_T^t \omega'(s)H(t,s)ds = -H(t,T)\omega(T) - \int_T^t \omega(s)\frac{\partial H(t,s)}{\partial s}ds$$

Therefore,

$$\int_T^t H(t,s) \int_a^b q(s,\xi) d\sigma(\xi)\, ds$$

$$\leq H(t,T)\omega(T) + \int_T^t \omega(s)\frac{\partial H(t,s)}{\partial s}\, ds - \int_T^t \frac{aH(t,s)g'(s,a)|\omega(s)|^{\frac{a+1}{a}}}{[r(s)]^{\frac{1}{a}}}\, ds \quad (6)$$

$$\leq H(t,T)\omega(T) + \int_T^t |h(t,s)\omega(s)|\, ds - \int_T^t \frac{aH(t,s)g'(s,a)|\omega(s)|^{\frac{a+1}{a}}}{[r(s)]^{\frac{1}{a}}}\, ds$$

Hence, according to Lemma 1, with

$$A = [aH(t,s)g'(s,a)]^{\frac{a}{a+1}}\frac{|\omega(s)|}{[r(s)]^{\frac{1}{a+1}}},$$

$$\lambda = \frac{a+1}{a}, \quad B = \left(\frac{a}{a+1}\right)^a |h(t,s)|^a \left\{\frac{r(s)}{[aH(t,s)g'(s,a)]^a}\right\}^{\frac{a}{a+1}},$$

We obtain that

$$|h(t,s)\omega(s)| - \frac{aH(t,s)g'(s,a)|\omega(s)|^{\frac{a+1}{a}}}{[r(s)]^{\frac{1}{a}}} \leq \frac{r(s)|h(t,s)|^{a+1}}{(a+1)^{a+1}[H(t,s)g'(s,a)]^a} \quad (7)$$

From (6), (7), we obtain that for all $t \geq T \geq t_0$,

$$\int_T^t \left[H(t,s)\int_a^b q(s,\xi)d\sigma(\xi) - \frac{r(s)|h(t,s)|^{a+1}}{(a+1)^{a+1}[H(t,s)g'(s,a)]^a}\right]ds \leq H(t,T)\omega(T) \leq H(t,t_0)\omega(T)$$

$$(8)$$

For $t \geq t_0$

$$\int_{t_0}^t \left[H(t,s)\int_a^b q(s,\xi)d\sigma(\xi) - \frac{r(s)|h(t,s)|^{a+1}}{(a+1)^{a+1}[H(t,s)g'(s,a)]^a}\right]ds$$

$$= \left\{\int_{t_0}^T + \int_T^t\right\}\left[H(t,s)\int_a^b q(s,\xi)d\sigma(\xi) - \frac{r(s)|h(t,s)|^{a+1}}{(a+1)^{a+1}[H(t,s)g'(s,a)]^a}\right]ds$$

$$\leq H(t,t_0)\int_{t_0}^T \int_a^b q(s,\xi)d\sigma(\xi)ds + H(t,t_0)\omega(T)$$

It follows that

$$\limsup_{t\to\infty}\frac{1}{H(t,t_0)}\int_{t_0}^t \left[H(t,s)\int_a^b q(s,\xi)d\sigma(\xi) - \frac{r(s)|h(t,s)|^{a+1}}{(a+1)^{a+1}[H(t,s)g'(s,a)]^a}\right]ds$$

$$\leq \int_{t_0}^T \int_a^b q(s,\xi)d\sigma(\xi)ds + \omega(T) < \infty$$

which contradicts (4). This completes the proof of Theorem 2.2.1.

Corollary 1 If the condition (4) is replaced by

$$\limsup_{t\to\infty}\frac{1}{H(t,t_0)}\int_{t_0}^{t}H(t,s)\int_{a}^{b}q(s,\xi)d\sigma(\xi)ds=\infty, \tag{9}$$

and

$$t\to\infty\;\frac{1}{H(t,t_0)}\int_{t_0}^{t}\frac{r(s)|h(t,s)|^{a+1}}{[H(t,s)g'(s,a)]^{a}}ds<\infty, \tag{10}$$

then Eq.(1) is oscillatory.

Example 1 Consider

$$[c\varphi_a(x')]'+\int_{1}^{2}e^{t+\xi}|x(t+\xi)|^{a}d\sigma(\xi)=0,\; t\geq 1 \tag{11}$$

where $t_0=1, a=1, b=2, q(t,\xi)=e^{t+\xi}, g(t,\xi)=t+\xi$. Let $H(t,s)=(t-s)^2$, then

$$\limsup_{t\to\infty}\frac{1}{H(t,t_0)}\int_{t_0}^{t}\left[H(t,s)\int_{a}^{b}q(s,\xi)d\sigma(\xi)-\frac{c|h(t,s)|^{a+1}}{(a+1)^{a+1}[H(t,s)g'(s,a)]^{a}}\right]ds$$

$$=\;\limsup_{t\to\infty}\frac{1}{(t-1)^2}\int_{1}^{t}\left[(t-s)^2\int_{1}^{2}e^{t+\xi}d\sigma(\xi)-\frac{c|2(s-t)|^{a+1}}{(a+1)^{a+1}(t-s)^{2a}}\right]ds$$

$$=\;\limsup_{t\to\infty}\frac{1}{(t-1)^2}\int_{1}^{t}\left[(e^2-e^1)(t-s)^2e^s-\frac{c2^{a+1}}{(a+1)^{a+1}(t-s)^{a-1}}\right]ds$$

$$=\;\infty$$

Hence, all solutions of Eq. (11) are oscillatory by Corollary 1.

Theorem 2 *Let functions H, h, ρ be the same as in Theorem* 1. *Moreover, suppose that*

$$0<\inf_{s\geq t_0}\left[\liminf_{t\to\infty}\frac{H(t,s)}{H(t,t_0)}\right]\leq\infty, \tag{12}$$

$$\limsup_{t\to\infty}\frac{1}{H(t,t_0)}\int_{t_0}^{t}\frac{r(s)|h(t,s)|^{a+1}}{[H(t,s)g'(s,a)]^{a}}ds<\infty. \tag{13}$$

If there exists a function $\varphi\in C([t_0,\infty);R)$, such that for every $T\geq t_0$,

$$\limsup_{t\to\infty}\frac{1}{H(t,T)}\int_{T}^{t}\left[H(t,s)\int_{a}^{b}q(s,\xi)d\sigma(\xi)-\frac{r(s)|h(t,s)|^{a+1}}{(a+1)^{a+1}[H(t,s)g'(s,a)]^{a}}\right]ds\geq\varphi(T),$$

$$\tag{14}$$

$$\int_{t_0}^{\infty} \frac{\varphi_+^{\frac{a}{a+1}}(s)g'(s,a)}{[r(s)]^{\frac{1}{a}}}ds = \infty, \tag{15}$$

where $\varphi_+(s) = \max\{\varphi(s), 0\}$, then Eq. (1) is oscillatory.

Proof Without loss of generality, we may assume that there exists a solution x(t) of Eq. (1) such that $x(t) \neq 0$, $t \geq t_0$. Define $\omega(t)$ as in (5). As in the proof of Theorem 1, we can have (6) and (8). Thus, we have

$$\limsup_{t\to\infty} \frac{1}{H(t,T)} \int_T^t \left[H(t,s) \int_a^b q(s,\xi)d\sigma(\xi) - \frac{r(s)|h(t,s)|^{a+1}}{(a+1)^{a+1}[H(t,s)g'(s,a)]^a} \right] ds \leq \omega(T)$$

for $t \geq T \geq t_0$. Thus by (14) we have

$$\varphi(T) \leq \omega(T), T \geq T_0, \tag{16}$$

$$\limsup_{t\to\infty} \frac{1}{H(t,T_0)} \int_{T_0}^t H(t,s) \int_a^b q(s,\xi)d\sigma(\xi)ds \geq \varphi(T_0). \tag{17}$$

Define

$$F(t) = \frac{1}{H(t,T_0)} \int_{T_0}^t |h(t,s)\omega(s)|ds,$$

$$G(t) = \frac{1}{H(t,T_0)} \int_{T_0}^t \frac{aH(t,s)g'(s,a)|\omega(s)|^{\frac{a+1}{a}}}{[r(s)]^{\frac{1}{a}}} ds,$$

by (6), we obtain that for all $t \geq T_0$,

$$G(t) - F(t) \leq \omega(T_0) - \frac{1}{H(t,T_0)} \int_{T_0}^t H(t,s) \int_a^b q(s,\xi)d\sigma(\xi)ds \tag{18}$$

Then, by (17) and (18), we see that

$$\liminf_{t\to\infty}[G(t) - F(t)]$$

$$\leq \omega(T_0) - \limsup_{t\to\infty} \frac{1}{H(t,T_0)} \int_{T_0}^t H(t,s) \int_a^b q(s,\xi)d\sigma(\xi)ds \leq \omega(T_0) - \varphi(T_0) < \infty. \tag{19}$$

Consider a sequence $\{T_n\}_{n=1}^{\infty} \in (t_0, \infty)$, $\lim_{n\to\infty} T_n = \infty$, and such that

$$\lim_{n\to\infty}[G(T_n) - F(T_n)] = \liminf_{t\to\infty}[G(t) - F(t)].$$

In view of (19), there exists a constant M, such that

$$G(T_n) - F(T_n) \leq M, \ n = 1, 2, \ldots \tag{20}$$

Now, we can prove that

$$\int_{T_0}^{\infty} \frac{|\omega(s)|^{\frac{a}{a+1}} g'(s,a)}{[r(s)]^{\frac{1}{a}}} \, ds < \infty. \tag{21}$$

If not, there exists a $T_1 > T_0$, such that $\int_{T_0}^{t} \frac{|\omega(s)|^{\frac{a}{a+1}} g'(s,a)}{[r(s)]^{\frac{1}{a}}} \, ds \geq \frac{\mu}{\alpha\eta}, t < T.$

for any positive number μ, where η is a positive number, and η satisfying

$$\inf_{s \geq t_0} \left[\liminf_{t \to \infty} \frac{H(t,s)}{H(t,t_0)} \right] > \eta > 0, \tag{22}$$

so for all $t > T_1$

$$\begin{aligned} G(t) &= \frac{a}{H(t,T_0)} \int_{T_0}^{t} H(t,s) d \left(\int_{T_0}^{s} \frac{|\omega(\tau)|^{\frac{a}{a+1}} g'(\tau,a)}{[r(\tau)]^{\frac{1}{a}}} \right) = -\frac{a}{H(t,T_0)} \int_{T_0}^{t} \frac{\partial H}{\partial s}(t,s) \left(\int_{T_0}^{s} \frac{|\omega(\tau)|^{\frac{a}{a+1}} g'(\tau,a)}{[r(\tau)]^{\frac{1}{a}}} \right) \\ &\geq -\frac{a}{H(t,T_0)} \int_{T_1}^{t} \frac{\partial H}{\partial s}(t,s) \left(\int_{T_0}^{s} \frac{|\omega(\tau)|^{\frac{a}{a+1}} g'(\tau,a)}{[r(\tau)]^{\frac{1}{a}}} \right) \geq -\frac{\mu}{\eta H(t,T_0)} \int_{T_1}^{t} \frac{\partial H}{\partial s}(t,s) ds = \frac{\mu H(t,T_1)}{\eta H(t,T_0)}. \end{aligned}$$

From (22), there exists a $T_2 \geq T_1$ such that $\frac{H(t,T_1)}{H(t,t_0)} \geq \eta$, and $G(t) \geq \mu$ for $t \geq T_2$. Since μ is an arbitrary constant, we conclude that $\lim\limits_{t \to \infty} G(t) = \infty$. Therefore

$$\lim_{n \to \infty} G(T_n) = \infty. \tag{23}$$

It follows from (20) and (23) that

$$\lim_{n \to \infty} F(T_n) = \infty. \tag{24}$$

Then we have, $\frac{F(T_n)}{G(T_n)} - 1 \geq -\frac{M}{G(T_n)} > -\frac{1}{2}$, for n large enough.
This and (24) imply that

$$\lim_{n \to \infty} \frac{F^{a+1}(T_n)}{G^a(T_n)} = \infty. \tag{25}$$

On the other hand, by Hölder inequality, we have

$$\begin{aligned} F(T_n) &= \frac{1}{H(T_n, T_0)} \int_{T_0}^{T_n} |h(T_n, s)\omega(s)| ds \\ &= \left\{ \frac{a}{H(T_n, T_0)} \int_{T_0}^{T_n} H(T_n, s) \frac{|\omega(s)|^{\frac{a}{a+1}} g'(s,a)}{[r(s)]^{\frac{1}{a}}} ds \right\}^{\frac{a}{a+1}} \times \left\{ \frac{1}{a^a H(T_n, T_0)} \int_{T_0}^{T_n} \frac{r(s)|h(T_n,s)|^{a+1}}{[H(T_n,s)g'(s,a)]^a} ds \right\}^{\frac{1}{a+1}} \\ &\leq \frac{G^{\frac{a}{a+1}}(T_n)}{a^{\frac{a}{a+1}}} \left\{ \frac{1}{H(T_n, T_0)} \int_{T_0}^{T_n} \frac{r(s)|h(T_n,s)|^{a+1}}{[H(T_n,s)g'(s,a)]^a} ds \right\}^{\frac{1}{a+1}} \end{aligned}$$

and therefore, $\dfrac{F^{a+1}(T_n)}{G^a(T_n)} \le \dfrac{1}{a^a H(T_n, t_0)} \displaystyle\int_{t_0}^{T_n} \dfrac{r(s)|h(T_n,s)|^{a+1}}{[H(T_n,s)g'(s,a)]^a}\, ds.$

It follows from (25) that $\displaystyle\lim_{n\to\infty} \dfrac{1}{H(T_n,t_0)} \int_{t_0}^{T_n} \dfrac{r(s)|h(T_n,s)|^{a+1}}{[H(T_n,s)g'(s,a)]^a}\, ds = \infty,$

and therefore $\displaystyle\lim_{t\to\infty} \dfrac{1}{H(t,t_0)} \int_{t_0}^{t} \dfrac{r(s)|h(t,s)|^{a+1}}{[H(t,s)g'(s,a)]^a}\, ds = \infty,$ which contradicts (13),

Hence, (21) holds. Then, it follows from (16), that $\displaystyle\int_{T_0}^{\infty} \dfrac{\varphi_+^{\frac{a}{a+1}}(s)g'(s,a)}{[r(s)]^{\frac{1}{a}}}\, ds \le$

$\displaystyle\int_{T_0}^{\infty} \dfrac{|\omega(s)|^{\frac{a}{a+1}}g'(s,a)}{[r(s)]^{\frac{1}{a}}}\, ds < \infty$

which contradicts (15). This completes the proof of Theorem 2.

3 Conclusion

In this paper, by using the Riccati transformation technique [6, 7], we have established some new oscillation criteria of second-order functional differential equations.

References

1. Yu YH, Fu XL (1991) Oscillation of second order nonlinear neutral equation with continuous distributed deviating argument. Rad Mat 7:167–176
2. Li HJ, Yeh CC (1995) Oscillations of half-linear second order differential equations. Hiroshima Math J 25:585–594
3. Hsu HB, Yeh CC (1996) Oscillation theorems for second order half-linear differential equations. Appl Math Lett 9:71–77
4. Manojlović (1999) Oscillation criteria for second-order half-linear differential equations. Math Comput Model 30:109–119
5. Wang QR (2001) Oscillation and asymptotic for second-order half-linear differential equations. Appl Math Comput 122:253–266
6. Yang XJ (2002) Oscillation results for second-order half-linear differential equations. Math Comput Model 36:503–507
7. Wang PG, Li XW (2003) Further results on oscillation of a class of second-order neutral equations. Comput Appl Math 157:407–418

An Outlier Detection Method for Robust Manifold Learning

Chun Du, Jixiang Sun, Shilin Zhou and Jingjing Zhao

Abstract Manifold learning algorithms have been widely used in data mining and pattern recognition. Despite their attractive properties, most manifold learning algorithms are not robust to outliers. In this paper, a novel outlier detection method for robust manifold learning is proposed. First, the contextual distance based reliability score is proposed to measure the likelihood of each sample to be a clean sample or an outlier. Second, we design an iterative scheme on the reliability score matrix to detect outliers. By considering both local and global manifold structure, the proposed method is more topologically stable than RPCA method. The proposed method can serve as a preprocessing procedure for manifold learning algorithms and make them more robust, as observed from our experimental results.

Keywords Robust manifold learning · Outlier detection · Contextual distance

1 Introduction

Over the past 10 years, many manifold learning methods have been proposed to discover the intrinsic geometric structures of high-dimensional data sets. Some representative ones, such as Locally Linear Embedding (LLE) [1], Isometric mapping (ISOMAP) [2], Local Tangent Space Alignment (LTSA) [3], Local Spline Embedding [4] and Maximal Linear Embedding [5] have been widely used in the areas of data mining and pattern recognition. Despite their attractive properties, manifold learning methods are often sensitive to outliers.

As defined by Hawkins [6], an outlier is an observation that deviates so much from other observations as to arouse suspicion that it was generated by a different

C. Du (✉) · J. Sun · S. Zhou · J. Zhao
School of Electronic Science and Engineering, National University of Defense Technology,
Changsha, Hunan, People's Republic of China
e-mail: yxduchun@gmail.com

mechanism. To deal with noisy data, researchers present many outlier detection schemes based on clustering, statistics, distance, depth and density [7]. Since the manifold assumption of local linear and global nonlinear is not taken into account, these traditional outlier detection schemes are not suitable for manifold learning. Recently, Chang and Yeung [8] proposed an outlier detection method based on the robust principal component analysis (RPCA). Their method first uses the weighted PCA to obtain the tangent space, and then detects outliers based on the projecting distances between each sample and its tangent space. RPCA based method has been successfully applied to extend LLE algorithm. However, when the intrinsic dimension of dataset is unknown, this method will fail to run. Moreover, the performance of outlier detection is sensitive to the algorithm parameters. These two disadvantages limit the applications of RPCA based method.

To improve the robustness of existing manifold learning algorithms, we proposed a novel outlier detection method based on contextual distance [9]. The contextual distance, which is a structural descriptor, can characterize the geometric structure better than the traditional Euclidean distance. Combining this method with LTSA algorithm, we obtain a robust version of manifold learning algorithm.

The paper is organized as follows. Section 2 gives a brief description of the contextual distance. Section 3 presents the proposed outlier detection method and gives a robust LTSA algorithm. Experimental results on Swiss roll and S curve data are described in Sect. 4. Section 5 is the conclusion of this paper.

2 Contextual Distance

In traditional outlier detection methods, Euclidean distance is widely used to measure the differences between two samples. And yet, in many practical applications, the data spaces are often high-dimensional, nonlinear and non-Euclidean. To depict the geometric differences between samples in non-Euclidean space, Zhao et al. [9] proposes the contextual distance.

Given a D-dimensional data set $X = \{x_1, x_2, \ldots, x_N\} \in R^{D \times N}$, sampled from a d-dimensional manifold $M(d < D)$. For each sample x_i, $i = 1, 2, \ldots, N$, we define its contextual set as its K nearest neighbors $S_i = \{x_i, x_{i1}, x_{i2}, \ldots, x_{iK}\}$, then the global structural characteristics of S_i can be defined as $f(S_i)$, where $f(\cdot)$ is a structural descriptor. As indicated in [9], if a neighbor sample x_{ij} complies with the structure of S_i, then removing x_{ij} from S_i will not affect the structure much. If x_{ij} is an outlier, then removing x_{ij} from S_i will change the structure significantly. Therefore, the contribution of x_{ij} to the integrity of the structure can be defined as follows:

$$\delta f_{ij} = \left| f(S_i) - f(S_i / x_{ij}) \right| \tag{1}$$

Then the contextual distance between x_i and x_{ij} is defined as:

$$d(x_i, x_j) = \left| \delta f_i - \delta f_{ij} \right| \tag{2}$$

3 Proposed Method

3.1 Weight Based on Coding Length

How to compute the reliability weight of each sample to be an outlier is still an open problem. The heat kernel weight, which is defined based on the Euclidean distance, is widely used in traditional outlier detection algorithms. And yet, since it cannot depict the geometric structure information between samples, the heat kernel weight is not suitable for the non-Euclidean data space. In this section, we use the coding length [10], which is one representative contextual distance, to obtain the reliability weight of each sample.

For each contextual set $S_i = \{x_i, x_{i1}, x_{i2}, \ldots, x_{iK}\}$, let its K nearest neighborhood matrix $X_i = [x_i, x_{i1}, x_{i2}, \ldots, x_{iK}] \in R^{D \times (K+1)}$, then the coding length of S_i can be defined as follows:

$$L(S_i) = \frac{K+1+N}{2} \log \det \left(I + \frac{N}{\varepsilon^2 (K+1)} \bar{X}_i \bar{X}_i^{\mathrm{T}} \right) + \frac{N}{2} \log \left(1 + \frac{\bar{x}_i^{\mathrm{T}} \bar{x}_i}{\varepsilon^2} \right) \qquad (3)$$

where $\det(\cdot)$ is the determinant operator, $\bar{x}_i = \frac{1}{K+1} X_i e$ is the mean vector of X_i, e is the $K+1$ dimensional all-one column vector, I is the $D \times D$ identity matrix, $\bar{X}_i = X_i - \bar{x}_i e^{\mathrm{T}}$ is the center matrix and ε is allowable distortion.

Based on Eq. (1), the contribution of x_{ij} to the integrity of the structure is:

$$\delta L_{ij} = \left| L(S_i) - L(S_i/x_{ij}) \right| \qquad (4)$$

Furthermore, the reliability weight of x_{ij} can be defined as [11]:

$$w_{ij} = \exp \left(- \frac{(\delta L_{ij} - \delta \bar{L}_i)^2}{2\sigma_i^2} \right) \qquad (5)$$

where $\delta \bar{L}_i$ and σ_i^2 are the mean and the standard deviation of $\{\delta L_{i0}, \delta L_{i1}, \delta L_{i2}, \ldots, \delta L_{iK}\}$.

3.2 Iterative Outlier Detection Scheme

For each local neighborhood, the reliability weight defined in Eq. (5) can characterize the likelihood of each sample to be an outlier. The smaller the value of w_{ij}, the more likely it is that x_{ij} is an outlier. Intuitively, one can set a threshold T and seek possible outliers by judging whether $w_{ij} < T$ or not. However, this scheme is not robust in many practical situations. First, the outliers based on local structure property may not be the true outliers in the sense of global structure. Furthermore, how to set a uniform threshold for all local neighborhoods is also a difficult task. In this section, we consider both local and global structure, and design an iterative scheme to detect outliers.

For each sample x_i, we set its reliability weight vector as $W^i = (W_1^i, W_2^i, \ldots W_N^i)^T$ whose elements are given below:

$$W_j^i = \begin{cases} w_{ij}, & \text{if } x_j \text{ belongs to the K nearest neighborhood of } x_i \\ 0, & \text{else} \end{cases} \quad (6)$$

Furthermore, we define the reliability weight matrix as $W = [W^1, W^2, \ldots, W^N]$. To improve the robustness of outlier detection algorithm, we consider all local neighborhoods involving x_i and define the average reliability score as follows:

$$r_i = \frac{1}{N_i} \sum_{j=1}^{N} W_i^j \quad (7)$$

where N_i is the non-zero element number of the ith row in matrix W.

Note that W_i^j denotes the contribution of x_i to the jth local neighborhood structure. Then r_i can be considered as the average contribution of x_i to the global structure. For a clean sample, since it comes from the true data structure, its average reliability score should be large, and it can be covered by considerable local neighborhoods. In contrast, for an outlier, since it deviates away from the data structure, its average reliability score is much small, and the number of local neighborhoods covering x_i is also not large. Motivated by it, we propose a novel criterion to distinguish clean samples and outlier samples as follows:

$$B(x_i) = \begin{cases} 1, & \text{if } r_i < \theta_r \text{ or } N_i < \theta_N \\ 0, & \text{else} \end{cases} \quad (8)$$

where $B(x_i)$ is an indicated function. $B(x_i) = 1$ means that x_i is an outlier and $B(x_i) = 0$ means x_i is a clean sample. Following the theory of robust statistics [12], we set the reliability score threshold θ_r and the neighborhood threshold θ_N as

$$\theta_r = \sum_{i=1}^{N} r_i/(2N), \; \theta_N = \sum_{i=1}^{N} N_i/(2N) \quad (9)$$

As depicted in Figs. 1c and 2c, we can detect the most of outliers by using Eq. (8) once. However, some outliers may still yet to be found, especially when partial outliers are close to each other. To solve this problem, we present an iterative outlier detection scheme as follows:

Input: Data set X, the initial parameter K
Output: Outliers set L

Step 1. Determine the $K - $ NN neighborhood $X_i = [x_i, x_{i1}, x_{i2}, \ldots, x_{iK}]$ for each sample x_i, compute the reliability weight based on coding length.
Step 2. Construct the reliability weight matrix W using Eqs. (5) and (6). Set $L = \emptyset$, $t = 0$.
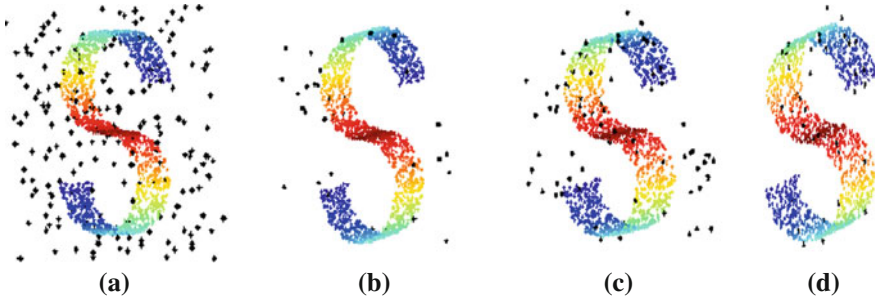
Fig. 1 The outlier detection results of S curve data set: **a** noisy data; **b** the result of the RPCA method; **c** the result of the first iteration; **d** the result of the last iteration

Step 3. Repeat the following steps:

1. $t = t + 1$.
2. Find the detected outliers in L and make the corresponding rows and columns of W be zero.
3. For each row of W, compute its average reliability score r_i and count the local neighborhood number N_i.
4. Update the reliability score threshold θ_r and the local neighborhood threshold θ_N using Eq. (9).
5. Detect new outliers using Eq. (8) and set the result as U_t, then let $L = L \cup U_t$.

Until the elements of L do not change anymore.

Considering the proposed method as a preprocessing step, we can obtain some robust versions of manifold learning algorithm. The low-dimensional embedding results can be easily solved by running these algorithms on the clean data set $X - L$. In this paper, we use our outlier detection method to improve the robustness of LTSA algorithm.

4 Experimental Results

4.1 Experiments of Outlier Detection

The test data includes S curve and Swiss roll, which are two representative manifold data sets. As shown in Figs. 1a and 2a, 1500 clean points are randomly sampled from the original manifold and 200 uniform distributed outliers are added as outliers (marked as '*'). To evaluate the performances of outlier detection, we apply our method and RPCA method to S curve and Swiss roll data, respectively. For RPCA, we set the neighborhood parameter $K = 15$ and set the uniform threshold $\theta_{\text{RPCA}} = 0.5$ like [8]. For ease of comparison, the neighborhood parameter K in our method is also set to 15.
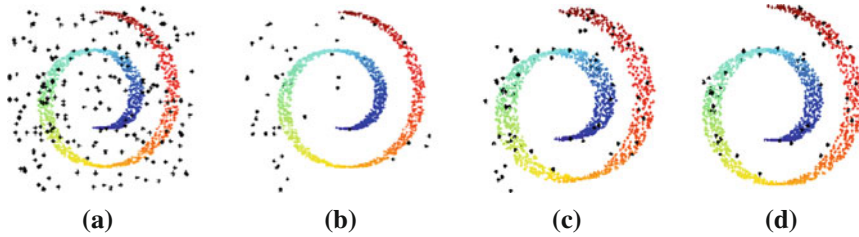
Fig. 2 The outlier detection results of Swiss roll data set: **a** noisy data; **b** the result of the RPCA method; **c** the result of the first iteration; **d** the result of the last iteration

The outlier detection results on S curve and Swiss roll data set are shown in Figs. 1 and 2. Shown in Fig. 1c is the first iterative result of our method. It can be clearly seen that many outliers are removed. However, the result of the first iteration is not encouraging and some outliers which are close to each other still exist. This is the reason why we design an iterative scheme to detect outliers. The final outlier detection result is shown in Fig. 1d. Comparing Fig. 1b with Fig. 1d, we find that our method is more topologically stable than the RPCA method. Obviously, RPCA method can remove the vast majority outliers. However, some outliers that deviate far away from the data structure still exist. These outliers may destroy the manifold structure badly. Although some outliers near to data structure remain in our method, they cannot change the manifold structure too much. Shown in Fig. 2 is the results on Swiss roll data, we can draw the same conclusions as the S curve experiment.

4.2 Experiments of Robust Manifold Leaning

Combining the proposed outlier detection method with LTSA, we obtain a robust LTSA algorithm. To illustrate the effectiveness of robust LTSA, we perform experiments on S curve and Swiss roll data. Shown in and Figs. 3c and 4c are the results of LTSA and robust LTSA respectively. Compared with LTSA, the embedding results of robust LTSA vary the color more smoothly. It means that our outlier detection plays an important role in manifold learning when the data set is corrupted by outliers.

5 Experimental Results

In this paper, we propose an outlier detection method for robust manifold learning. To measure the likelihood of each sample to be a clean sample or an outlier, we present a contextual distance based reliability score. Based on the reliability score,
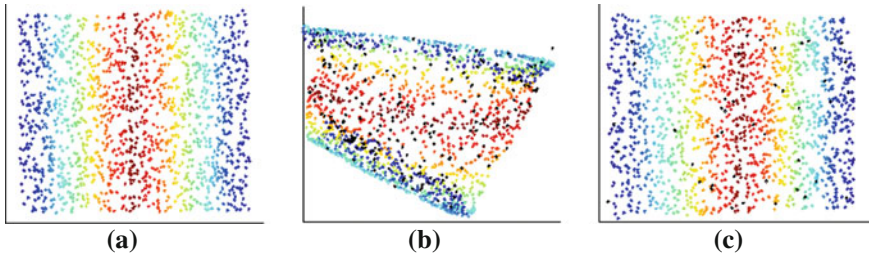
Fig. 3 Embedding results on S curve data: **a** by applying LTSA on clean data; **b** by applying LTSA on noisy data; **c** by applying robust LTSA on noisy data
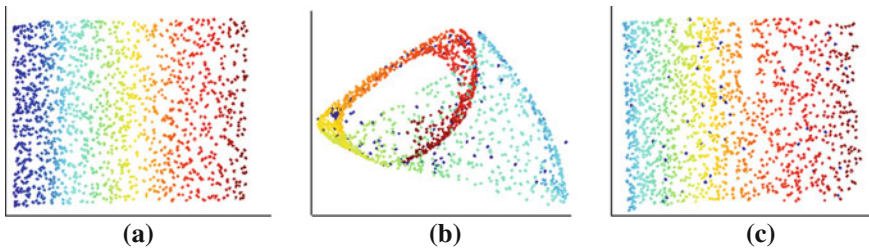
Fig. 4 Embedding results on Swiss roll data: **a** by applying LTSA on clean data; **b** by applying LTSA on noisy data; **c** by applying robust LTSA on noisy data

we design an iterative scheme to find the outlier set. The proposed method can serve as a preprocessing procedure for LTSA to form a robust version of LTSA. Experiments on several data sets such as S curve and Swiss roll data, verified the effectiveness of our proposed method. It should be pointed out that the proposed outlier detection method can be also used to extend other manifold learning algorithms and make them more robust. This is a potential direction of our future research.

References

1. Roweis ST, Saul LK (2000) Nonlinear dimensionality reduction by locally linear embedding. Science 290:2323–2326
2. Tenenbaum JB, Silva VD, Langford JC (2000) A global geometric framework for nonlinear dimensionality reduction. Science 290:2319–2323
3. Zhang Z, Zha H (2005) Principal manifolds and nonlinear dimension reduction via local tangent space alignment. SIAM J Sci Comput 26:313–338
4. Xiang S, Nie F, Zhang C (2009) Nonlinear dimensionality reduction with local spline embedding. IEEE Trans Knowl Data En 21(9):1285–1298
5. Wang R, Shan S, Chen X, Chen J (2011) Maximal linear embedding for dimensionality reduction. IEEE Trans Pattern Anal Mach Intell 33(9):1776–1792
6. Hawkins D (1980) Identification of outliers. Chapman and Hall, London

7. Maimon O, Rokach L (2005) Data mining and knowledge discovery handbook. Kluwer Academic Publishers, Dordrecht
8. Chang H, Yeung DY (2006) Robust locally linear embedding. Pattern Recogn 39:1053–1065
9. Zhao D, Lin Z, Tang X (2007) Contextual distance for data perception. In: Proceedings of international conference on computer vision, IEEE Computer Society Press, New York, pp 1–8
10. Ma Y, Derksen H, Hong W (2007) Segmentation of multivariate mixed data via lossy data coding and compression. IEEE Trans Pattern Anal Mach Intell 29(9):1546–1562
11. Zhao D, Lin Z, Tang X (2007) Laplacian PCA and its applications. In: Proceedings of international conference on computer vision, IEEE Computer Society Press, New York, pp 1–8
12. Huber PJ, Ronchetti EM (2009) Robust statistics, 2nd edn. Wiley, New York

Motif Identification Based on Local Structure Clustering

Junying Zhang and Yuling Xue

Abstract Network motif identification is significant in that motifs generally reflect functionalities. However, the task is greatly challenging due to diverse patterns of motifs possibly existent in a network and high computation complexity for large scale networks. In this study, we propose a network motif identification method, FCMD, based on clustering subgraphs according to their local structures. By modeling local-topological feature of a network with a feature vector, the approach maps motif identification problem into a clustering problem of feature vectors in a feature space, which greatly reduces computation complexity hence facilitates very large scale networks. Experiments on 8 real networks, including biochemical network, neural network, electronics circuit network, et al., indicates that the proposed method is very effective in motif detection with computation complexity approximately independent to the scale of the network.

Keywords Network motif identification · Feature vector · Clustering

1 Introduction

A lot of systems in nature demonstrate complex network features [1]. One of a significant feature is network motif, which refers to patterns of interconnections occurring in complex networks at numbers that are significantly higher than those in randomized networks [2–7]. These frequently occurred motifs generally demonstrate some functions. Hence identification of motifs from a network is a significant task to further understanding the functionalities of the network.

J. Zhang (✉) · Y. Xue
School of Computer Science and Technology, Xidian University,
Xi'an, People's Republic of China
e-mail: jyzhang@mail.xidian.edu.cn

The pioneer motif identification algorithms are Exhaustively-Enumerating algorithm (EEA) [3]. It acts by exhaustively enumerating all subgraphs of a given size in the network hence the runtime of such algorithm increases strongly with network size. Starting with networks where the interactions between nodes are represented by directed edges, EEA compares the real network to randomized networks with the same degree distribution as that of the network and only select patterns appearing in the real network at numbers significantly higher than those in the randomized networks with P value of less than 0.01. The approach is only suited for identification of small sized motif (e.g., the size of 4) but with an extensive computation for large network since it needs scan not only all possible patterns of size 4 over the network where motifs are to be identified, but requires generating a lot of random networks and scan these patterns over each of the random networks for understanding significance of the occurrence of each motif pattern in the network.

Several improvement algorithms are used to faster motif detection process, in which sampling algorithm (ESA) [6], the algorithm in [8] and the one in [9] are typical examples. All these improve computation, but can not guarantee globally optimal solutions.

This study proposes a novel local-topological feature clustering motif detection algorithm (FCMD) for motif identification. In the approach, we extract topological feature of a subgraph centered/defined by a node in the network, map it to a topological feature space, and cluster samples in the space into clusters, where each sample corresponds to a subgraph in the network. Then samples belonging to a same cluster have similar local topology, and a motif is detected by some post-processing procedure on each cluster. The distinctive characteristics of the approach are no need for generating random networks, no need to consider all possible patterns of size m exhaustively, and no need to scan the patterns throughout the network, hence drastically decreasing complexity on both computation and storage. Additionally, not like EEA/ESA, which can detect motifs of only size m (m is generally small, e.g., 3 or 4), the approach requires a user parameter k specifying local region of a node as a subgraph, and can detect motifs of different sizes simultaneously with the motif size generally much larger than that by EEA/ESA.

2 Method

Motif is the subgraph in a network which often occurs in the network. At first, it is the subgraph or local topology of the network; secondly, such topology appears in the network many times. Accordingly we represent a topology with a feature vector, and see if such topology/feature vector appears in the network many times. For the former, we provide a methodology to extract features of a topology with a feature vector in a feature space, and for the latter, we cluster all feature vectors in
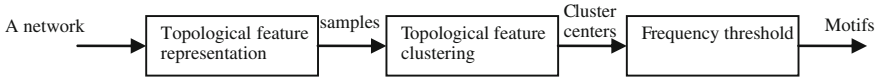
Fig. 1 The schematic diagram of the FCMD for motif identification

feature space to clusters, such that the number of appearance of a topology is reflected by size of a cluster that the topology belongs to. The schematic diagram is given in Fig. 1.

2.1 Topologic Feature Representation

Topology of a subgraph centered at a node is reflected by the node and connections within the subgraph.

For node i, its l-hops output set, denoted by $T_{l+}(i)$, is a set of nodes that the number of directed hops from node i to any node in $T_{l+}(i)$ is l; its l-hops input set, denoted by $T_{l-}(i)$, is a set of nodes that the number of directed hops from any node in $T_{l-}(i)$ to node i is l. The join of $T_{l+}(i)$ and $T_{l-}(i)$ is denoted by $T_l(i)$ being the set of nodes whose hop distance from node i is l.

For node i one can enumerate edges from the s-hops output/input set to l-hops output/input set, and denote them as $t_{s+l+}(i)$, $t_{s+l-}(i)$, $t_{s-l+}(i)$ and $t_{s-l-}(i)$ respectively.

Topology of a network at around node i can be represented by following factors:

1. The structural feature between node i and $T_l(i)$, including the number of nodes included in $T_{l+}(i)$, that included in $T_{l-}(i)$, the number of edges among nodes in $T_{l+}(i)$ and that among nodes in $T_{l-}(i)$, and the number of edges between nodes in $T_{l+}(i)$ and nodes in $T_{l-}(i)$. By denoting enumeration number being a respective t, these can be represented by a 6-dimensional vector of

$$v_l(i) = [t_{l+}(i), t_{l-}(i), t_{l+l+}(i), t_{l-l-}(i), t_{l+l-}(i), t_{l-l+}(i)] \qquad (1)$$

2. The structural feature between node set $T_s(i)$ and node set $T_l(i)$, including the number of edges from $T_s(i)$ to $T_l(i)$, and that from $T_l(i)$ to $T_s(i)$, which is represented by an 8-dimensional vector

$$v_{sl}(i) = [t_{s+l+}(i), t_{s-l-}(i), t_{s+l-}(i), t_{s-l+}(i), t_{l+s+}(i), t_{l-s-}(i), t_{l+s-}(i), t_{l-s+}(i)] \quad (2)$$

To study the local-topological structure of a graph, we refer a subgraph in this study as a k-region of node i: the k-region of node i is the subgraph in which all the nodes in the original network whose shortest path length (number of hops) from node i, or to node i, is less than or equal to k, where k is referred to as a user predefined radius.

Hence, for network G, the feature vector characterizing topology of the k-region of node i of the network G can be represented by a topological feature vector $v(i; k)$, which is

$$v(i; k) = [v_1(i), v_2(i), \ldots, v_k(i), v_{12}(i), v_{13}(i), \ldots, v_{(k-1)k}(i)]^T \qquad (3)$$

The dimensionality of the feature space is $d_k = 2k(2k+1)$ for $k \geq 3$, $d_k = 6$ and 20 for $k = 1$ and $k = 2$ respectively. As is known, for a graph or a subgraph, only adjacency matrix or its related form can represent its topological characteristic. The feature vector proposed above in fact can not wholly but can roughly represent the characteristic of the topology of the network or subgraph.

2.2 Clustering Features in Topological Feature Space

For a network G with n nodes and a fixed radius k, one can obtain n samples in topological feature space, each corresponding to the topological feature vector of the k-region of a node in the network.

Motif is the subgraph in a network which often occurs in the network. This means that they are common or similar in topological feature space, and hence they should be clustered into a same cluster. This is why we propose using clustering algorithm to cluster feature vectors for motif identification. We propose using affinity propagation (AP) clustering algorithm [10] for such clustering, which merits motif identification in the following aspects: (1) it is unnecessary to predetermine the number of clusters, which is just the case of motif identification; (2) the prototypical sample of each cluster obtained from the algorithm is some sample of the cluster rather than some pseudo sample (e.g., mean) computed from the samples in the cluster, reflecting the topological structure of a k-region of the corresponding node; (3) there is no need for symmetry of similarity metrics, which is just the case of motifs where the graph is allowed to be directed rather than undirected.

The user parameters in AP are set be $p =$ the medium value of similarities of all pair of samples, and $\lambda = 0.9$. The maximum iteration number is set to 1,000.

2.3 Motif Detection

Intuitively, the k-regions of all nodes in each cluster have their topology similar to each other, and we can conclude that the k-region of the prototypical node of a cluster is a motif which appears in the network with the number of times being the size of the cluster. However, notice that the k-region of a node and that of its neighborhood node generally have very similar topology, and the situation will be more serious for large k. Such appearance can not be enumerated twice. Hence we

propose a post-processing procedure for motif identification. The criterion of the procedure is: samples of the same cluster with common nodes are enumerated only once. Accordingly, the cluster with enumeration number larger than a threshold is considered being a motif. We set this threshold to be 4 in our experiment.

3 Experiments and Results

We performed experiments on 8 typical directed and weighted networks with our proposed algorithm, and compared the result obtained by the typical EEA algorithm in motif identification performance, focusing on if the precise motifs detected by the EEA are successfully included in the rough motifs detected by the FCMD.

The 8 typical directed and weighted networks in our experiments include gene regulatory network, neural connection network, biological signal conduction network and logic circuit network [11]. The number of nodes ranges from 122 to 512 and the number of edges ranges from 83 to 2,170. These networks are very typical and representative to a diverse of real networks. In our experiment, self-circles and connection weights are not in consideration.

We uses FCMD with the setting of $k = 1, 2, 3, 4$ respectively and EEA with the setting of $m = 3, 4$ for motif identification and comparison. Results are given in Table 1. It is seen that for even small setting of k, i.e., $k = 1$, all motifs detected by EEA are successfully detected by FCMD for all the networks except for C.elegans[1], though the FCMD is a rough motif identification algorithm. This indicates that the proposed FCMD is very effective in detection of precise motifs.

In fact, the number of detected rough motifs is generally larger than that of precise motifs. For example, for E. coli network, the number of precise motifs of size 3 by EEA is one, while that of the rough motifs by FCMD with $k = 1$ is 7. Additionally, the precise motif obtained by EEA is in fact included in the 7 rough motifs detected by the FCMD.

On the other hand, by noticing that the radius parameter k specifies how local the motifs are to be detected, the number of successfully detected motifs by the FCMD is decreasing as k increases for some of the networks. This is reasonable, since when one sets a larger k, he/she desires to detect larger motifs, which may appear in some networks, e.g., in the network of S838, or may appear but failed to

Table 1 The percentage of motifs successfully discovered by FCMD over the ones discovered by EEA

FCMD	E. coli	C.elegans1	C.elegans2	Humans	SeaUrchin	s208	s420	S838
K = 1	1	0.8571	1	1	1	1	1	1
K = 2	1	0.2857	0.1667	1	1	1	0.9545	1
K = 3	1	0.1429	0	1	1	0.9362	0.9545	1
K = 4	1	0.1429	0	1	1	0.9362	0.8636	1

be detected by the EEA with the setting of $m = 3, 4$, or may not appear in some networks, e.g., in the network of s420.

We have only one exception that there always exist some precise motifs detected by EEA fails to be detected by FCMD for $k = 1, 2, 3, 4$ and the exception is C.elegans[1]. The reason is that the detection scheme is different from that used in EEA and in FCMD. In EEA, motif is a subgraph defined by nodes, while in FCMD, motif is a subgraph in the region of a node (or centered at a node). We believe that a network might function from both types of motifs.

4 Conclusions

This paper proposes a methodology and algorithm for detecting motifs based on local-topological feature similarities at the regions of different nodes. The algorithm uses local-topological feature similarities and clusters the features for motif detection. The algorithm is very fast compared with the famous EEA algorithm, and is comparatively effective in detection of motifs. Besides, it is allowed to find more than one-sized motifs simultaneously.

References

1. Itzkovita S, Alon U (2005) Subgraphs and network motifs in geometric networks. Phys Rev E 71:026117
2. Shen-Orr SS, Milo R, Mangan S, Alon U (2002) Network motifs in the transcriptional regulation network of Escherichia coli. Nat Genet 31(1):64–68
3. Milo R, Shen-Orr S, Itzkovita S, Kashtan N, Chklovskii D, Alon U (2002) Network motifs: simple building blocks of complex networks. Science 298:10
4. Kashtan N, Itzkovita S, Milo R, Alon U (2004) Efficient sampling algorithm for estimating subgraph concentrations and detecting network motifs. Bioinformatics 20:1746–1758
5. Kashtan N, Itzkovita S, Milo R, Alon U (2004) Topological generalizations of network motifs. Phys Rev E 70:031909
6. Kashtan N, Itzkovitz S, Milo R, Alon U (2004) Efficient sampling algorithm for estimating subgraph concentrations and detecting network motifs. Bioinformatics 20(11):1746–1758
7. Artzy-Randrup Y et al (2004) Comment on "Network motifs: simple building blocks of complex networks" and "Superfamilies of evolved and designed networks". Science 305:1107c
8. Wernicke S (2006) Efficient detection of network motifs. IEEE/ACM Trans Comput Biol Bioinf 3(4):347–359
9. Barabasi A-L, Albert R (1999) Emergence of scaling in random network. Science 286:509–512

10. Itzhack R, Mogilevski Y, Louzoun Y (2007) An optimal algorithm for counting network motifs. Phys A 381:482–490
11. Grochow JA, Kellis M (2007) Network Motif discovery using subgraph enumeration and symmetry-breaking. In: Speed T, Huang H (eds) RECOMB 2007, vol 4453. Springer, Heidelberg, pp 92–106
12. Babu MM et al (2004) Structure and evolution of transcriptional regulatory networks. Struct Biol 14:283–291
13. Frey BJ, Dueck D (2007) Clustering by passing messages between data points. Science 315:972–976
14. http://www.weizmann.ac.il/mcb/UriAlon/

Computation of Synchronic Distance in Marked S-net

Li–li Wang, Xian-wen Fang and Dao-hao Liu

Abstract Synchronic distance is an important analyzing metric to describe the synchronic relationship between two events. Due to the difficulties to calculate the synchronic distance, which is determined by both the structure characteristics and the initial marking of the net system, no simple and feasible algorithms have been presented to get the synchronic distance of normal Petri nets. However, the computations of synchronic distance in some special subclass of Petri nets, such as marked T-graph, marked S-graph, and marked T-net, are relatively simple and feasible. The solution of the synchronic distance in another subclass of Petri net-marked S-net, is given in this paper. The synchronic distance of the marked S-net is directly obtained according to the structure of the net systems and the distribution situation of initial marking, which makes the solution of the synchronic distance become more feasible. The corresponding solution theorems are also given in the paper.

Keywords Synchronic distance · Marked S-net · Source transition · Sink transition · Behavior equivalence

1 Introduction

As a model for the description of the asynchronous and concurrent phenomenon systems Petri nets have shown its unique advantage, and have been widely used in many applications in various fields [1, 2], especially in concurrent system. The synchronization between information in the information system are very important, such as the synchronization between information sending, transferring, and

L. Wang (✉) · X. Fang · D. Liu
College of Mathematics, Anhui University of Science and Technology,
Huainan 232001, People's Republic of China
e-mail: wanglili198301@163.com

receiving in a information system, and the synchronization of the information within and between the media streams in the multimedia system [3, 4]. Base on the synchrony theory, the synchronic distance is typically used to describe the synchronization in the above actual systems. The synchronic distance is not only an analyzing metric to describe the synchronic relationship between two events, but also a tool to analyze the dynamic behavior of the systems. It is very suitable for measuring the maximum directional signal scale in the source model based on Petri nets. Synchronizer of Petri nets is currently the most effective modeling method of workflow, and has a wide applications in workflow, web service and other fields [5–8].

The concept of synchronic distance was first introduced by C.A. Petri in the article [9]. Because computation of synchronic distance involves both structure and initial marking of net systems, it needs to simulate the operation of the systems, which adds the difficulties to solve synchronous distance. However, the computations of synchronic distance in some special subclass of Petri nets such as marked T-graph [10], marked S-graph [11], and marked T-net [12], are relatively simple and feasible so far.

This paper presents an algorithm to calculate the synchronic distance of another subclass of Petri net-marked S-net. The synchronic distance of the marked S-net is directly obtained according to the structure of the net systems and the distribution situation of initial marking without simulating the operation of the net systems, as well as the marked T-graph and the marked S-graph. The presented method provides a more feasible way to calculate the synchronic distance.

2 Basic Concept

The basic concepts and conclusions of Petri nets can be found in details in the Ref. [11]. For the convenience of the later discussion, the related basic concepts, terminology, and notations are introduced in this section simply in this section.

Generally, a Petri net is denoted by a 4-tuple $\sum = (S, T; F, M_0)$, in which $N = (S, T; F)$ denotes a net, M_0 is an initial marking. M is said to be reachable from M_0 if there exists $\sigma \in T^*$ such that $M_0[\sigma > M. R(M_0)$ denotes the set of all possible markings reachable from M_0.

Let $N = (S, T; F)$ be a net, N is said to be a S-graph if $\forall t \in T: |{}^\bullet t| = |t^\bullet| = 1$. $\sum = (N, M_0)$ is said to be a marked S-graph if $N = (S, T; F)$ is a S-graph [11].

Let $\sum$ be a marked S-graph. For $x \in S \cup T$, the strongly connected components containing x are denoted as $SC(x)$. For $s \in S$, the maximum markings obtained by the place s during the implementation of the net $\sum$ are denoted as $M_{\max}(s)$, i.e. $M_{\max}(s) = \max\{M(s)|M \in R(M_0)\}$.

Lemma 1 [11] *Let* $\sum = (S, T; F, M_0)$ *be a marked S-graph. If $s_1 \in S$, then*

$$M_{\max}(s_1) = M_0(s_1) + \sum_{(s,s_1) \in F^+} M_0(s) \tag{1}$$

Theorem 1 [11] *Let* $\sum = (S, T; F, M_0)$ *be a marked S-graph ($\sum$ is also a proper net system). If $t_1, t_2 \in T$, the synchronic distance between t_1 and t_2 is divided into the following three cases.*

1. *If all the directed circuits which pass through t_1 also pass through t_2, and all the directed circuits which pass through t_2 also pass through t_1, then*

$$\sigma(t_1, t_2) = M_{\max}(^\bullet t_1) = M_{\max}(^\bullet t_2) \tag{2}$$

2. *If either of t_1 and t_2 are not on a directed circuit, then*

$$\sigma(t_1, t_2) = \max\{M_{\max}(^\bullet t_1), M_{\max}(^\bullet t_2)\} \tag{3}$$

3. *If there is a directed circuit C which passes through only one of t_1 and t_2, then $\sigma(t_1, t_2) = \infty$.*

   In which $M_{\max}(^\bullet t_i)\,(i = \{1, 2\})$ in Eqs. (2) and (3) can calculated by Eq. (1).

Definition 1 [11] Let $N = (S, T; F)$ be a net, N is called a S-net if $\forall t \in T$: $|^\bullet t| \leq 1$ and $|t^\bullet| \leq 1$.

$\sum = (N, M_0)$ is called a marked S-net if $N = (S, T; F)$ is a S-net.

Properties of S-net can be found in references [13] and [14]. Relevant definitions are introduced in the following before introducing the solution of synchronic distance in these net systems.

Definition 2 Let $N = (S, T; F)$ be a net, t is called a source transition of N if $^\bullet t = \phi$; t is said to be a sink transition of N if $t^\bullet = \phi$.

Definition 3 Let $\sum = (S, T; F, M_0)$ be a Petri net, $\forall x \in S \cup T$

1. $\Pr eS - Set(x) = \{y \in S | (y, x) \in F^*, x \in S \cup T\}$, in which $F^* = \{x, x\} \cup F^+$ (i.e., the set of places in all the paths of which the terminal points are x), $\Pr eS - Set(x)$ is called the set of preorder places of x. Obviously, if x is a place, then $x \in \Pr eS - Set(x)$.
2. $PostS - Set(x) = \{y \in S | (x, y) \in F^*, x \in S \cup T\}$, in which $F^* = \{x, x\} \cup F^+$ (i.e., the set of places in all the paths of which the starting points are x), $PostS - Set(x)$ is called the set of postorder place of x. Obviously, if x is a place, then $x \in PostS - Set(x)$.
3. $\Pr eT - Set(x) = \{y \in T | (y, x) \in F^*, x \in S \cup T\}$, in which $F^* = \{x, x\} \cup F^+$ (i.e., the set of transitions in all the paths of which the terminal points are x), $\Pr eT - Set(x)$ is called the set of preorder transition of x. Obviously, if x is a transition, then $x \in \Pr eT - Set(x)$.
4. $PostT - Set(x) = \{y \in T | (x, y) \in F^*, x \in S \cup T\}$, in which $F^* = \{x, x\} \cup F^+$ (i.e., the set of transitions in all the paths of which the starting points are x),

PostT − Set(x) is called the set of postorder transition of x. Obviously, if x is a transition, then $x \in PostT − Set(x)$.

From the above definitions, it can be defined in the following, if $\forall X \subseteq S(or \forall X \subseteq T)$:

1. $\Pr eS − Set(X) = \bigcup_{x \in X} \Pr eS − Set(x);$
2. $PostS − Set(X) = \bigcup_{x \in X} PostS − Set(x);$
3. $\Pr eT − Set(X) = \bigcup_{x \in X} \Pr eT − Set(x);$
4. $PostT − Set(X) = \bigcup_{x \in X} PostT − Set(x).$

Definition 4 [11] Let $\sum = (S, T; F, M_0)$ and $\sum' = (S', T; F', M_0')$ be two Petri nets, $\sum'$ and $\sum$ are said to be behavioral equivalent if there exists a surjective $f : R(M_0') \rightarrow R(M_0)$ such that (1) $f(M_0') = M_0$ (2) if $f(M') = M$, for $\forall \sigma \in T^*$: exist $M'[\sigma >$ in net $\sum'$ iff exist $M[\sigma >$ in net $\sum$.

3 Solution of Synchronic Distance in Marked S-net

When we discuss the synchronic distance between two transitions in a net system, it is generally assumed that $\sum$ is a proper net system, i.e., $\forall t \in T, \exists M \in R(M_0) : M[t >$. It easy to see that a proper system net $\sum$ is L1-live. Thus, it is assumed that net systems discussed in this paper are all proper systems.

The synchronic distance between transition t_i and transition t_j is denoted as $sd(t_i, t_j)$.

Two signs are introduced before giving the solution of the synchronic distance in S-net: for any given S-net $\sum = (S, T; F, M_0)$, the set of all the source transitions in net $\sum$ is denoted as $T_{source} = \{t | t \in T \wedge |^\bullet t| = 0\}$. The set of all sink transitions in net $\sum$ is denoted as $T_{\sin k} = \{t | t \in T \wedge |t^\bullet| = 0\}$.

Theorem 2 *Let $\sum = (S, T; F, M_0)$ be a marked S-net. If $\forall t \in T : |^\bullet t| = 1$ and $|t^\bullet| \leq 1, t_1, t_2 \in T$, then the computation method of the synchronic distance between transition t_1 and t_2 are the same as that of the marked S-graph.*

Proof For any given S-net $\sum$ which contains sink transitions while without source transition (i.e., $T_{source} = \phi, T_{\sin k} \neq \phi$), the S-net containing sink transitions can be converted into a S-net $\sum'$ without sink transitions by adding each output place to each sink transition (obviously the new output place does not affect behavior of the net system), under the condition of keeping the behavior characteristic of each transition unchanged. Thus, the S-net is converted into a behavior equivalent S-graph. Therefore, the solution of synchronic distance in a S-net system only having sink transitions can be converted into the solution of the S-graph.   □

Any given S-net containing sink transitions can all be converted into S-net which does not have sink transitions when the synchronic distances of each pair of transitions are kept to be constant by adding a output place for each sink transition. Therefore, it is assumed that S-net in the following theorem does not contain sink transitions in this paper.

Theorem 3 *Let* $\sum = (S, T; F, M_0)$ *be a marked S-net. If* $T_{source} \neq \phi$, *then for each pair of* $t_1, t_2 \in T$, *synchronic distance can divided into the following two cases*:

1. *For* $\forall t_i \in PostT - Set(T_{source}), t_j \in T{:}sd(t_i, t_j) = \infty$
2. *For* $\forall t_i, t_j \notin PostT - Set(T_{source})$: *the solution of* $sd(t_i, t_j)$ *are the same as that of the marked S-graph (see Theorem 1).*

Proof For (1): we proof it from two cases:

1. If $t_i \in PostT - Set(T_{source}), t_j \notin PostT - Set(T_{source})$, because $t_i \in PostT - Set(T_{source})$, by firing $\forall t' \in \Pr eT - Set(t_i) \cap T_{source}$ any number of times, such that $t'^\bullet$ get any number of markings, which can flow to $^\bullet t_i$ finally by firing series of transitions. So $\exists M \in R(M_0)$, for any positive integer k, $\exists \sigma_1, M[\sigma_1 >$ such that $^\bullet t_i$ gets k_1 (which $k_1 > k$) markings, it exists $M[\sigma_1 t_i^{k_1} > (denotes \sigma_2 = \sigma_1 t_i^{k_1})$ satisfying $\#(t_j/\sigma_2) = 0, \#(t_i/\sigma_2) > k$, which indicates that t_i and t_j are not in weak fair relation, therefore t_i and t_j also are not in meta-fair relation. So $sd(t_i, t_j) = \infty$.
2. If $t_i, t_j \in PostT - Set(T_{source})$, It is assumed that there existing a directed path from t_i to t_j, for $\forall M \in R(M_0)$, by firing transitions any number of times from $\forall t' \in \Pr eT - Set(t_i) \cap T_{source}$ to t_i such that $^\bullet t_i$ gets any number of markings, thus t_i can fire any number of times and t_j fires zero times. Therefore, it is easy to know that t_i and t_j are not in weak fair relation, so they are also not in meta-fair relation. So $sd(t_i, t_j) = \infty$. $\square$

(1) and (2) shows that the conclusion in (1) is correct.

For (2): If $\forall t_i, t_j \notin PostT - Set(T_{source})$, we know that transitions in T_{source} firing does not impact the dynamic behavior of t_i and t_j, so when analyzing synchronic distance between t_i and t_j, we can ignore transitions in $PostT - Set(T_{source})$ firing. Therefore for this case we can compute synchronic distance between t_i and t_j by adopting the methods of S-graph directly.

4 Example

Example: Figure 1 shows a marked S-net$\sum_1$ which contains source transitions and sink transitions. We will give synchronic distance between each pair of transitions. From Fig. 1 we can know that net $\sum_1$ have a source transition t_8 and

Fig. 1 A marked S-net$\sum_1$ which contains source transitions and sink transitions
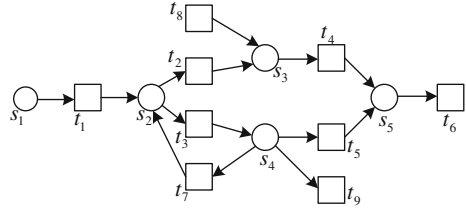
two sink transitions t_6, t_9, So $T_{source} = t_8, T_{sink} = \{t_6, t_9\}$. We can obtain $PostT - Set(T_{source}) = \{t_4, t_6, t_8\}$ by source transitions. In the following we apply above Theorem 2 and Theorem 3 to calculate synchronic distance between each pair of transitions:

1. t_4, t_6, t_8 are postorder transitions of source transitions respectively, therefore:

$$sd(t_4, t_i) = \infty \quad i = 1, 2, 3, 5, 6, 7, 8, 9$$
$$sd(t_6, t_i) = \infty \quad i = 1, 2, 3, 4, 5, 7, 8, 9$$
$$sd(t_8, t_i) = \infty \quad i = 1, 2, 3, 5, 6, 7, 9$$

2. Remaining transitions are all not postorder transitions of source transitions, from Theorem 2 it can know that sink transitions are not considered particularly when calculating the synchronic distance. So we get the synchronic distance between these remaining pair of transitions are the same as that of the marked S-graph.

Both t_3 and t_7 are in the directed circuit, but others are not in any directed circuits. Therefore:

$$t_1: sd(t_1, t_2) = \max\{M_{\max}(^\bullet t_1), \; M_{\max}(^\bullet t_2)\} = 3$$
$$sd(t_1, t_3) = \infty$$
$$sd(t_1, t_5) = \max\{M_{\max}(^\bullet t_1), \; M_{\max}(^\bullet t_5)\} = 4$$
$$sd(t_1, t_7) = \infty$$
$$sd(t_1, t_9) = \max\{M_{\max}(^\bullet t_1), \; M_{\max}(^\bullet t_9)\} = 4$$
$$t_3: sd(t_3, t_5) = \infty$$
$$sd(t_3, t_7) = M_{\max}(^\bullet t_3) = M_{\max}(^\bullet t_7) = 4$$
$$sd(t_3, t_9) = \infty$$
$$t_5: sd(t_5, t_7) = \infty$$
$$sd(t_5, t_9) = \max\{M_{\max}(^\bullet t_5), \; M_{\max}(^\bullet t_9)\} = 4$$
$$sd(t_3, t_9) = \infty$$

5 Conclusion

A subclass of Petri net—marked S-net is presented in this paper, in which the solution of synchronic distance is also presented simply and feasible. Synchronic distance of marked S-net is directly calculated according to the structure of net systems and the distributions situation of initial marking, just as the calculation of the marked S-graph and the marked T-graph. Corresponding solution theorem of the synchronic distance is then derived. Examples are finally illustrated to further explain the method.

References

1. Yuan C (2005) The principle of Petri nets. Publishing House of Electronics Industry, Beijing
2. Zhao W, Huang Y, Yuan C (2008) Synchronic distance based workflow logic specification. In: Proceedings of the 2008 10th IEEE international conference on high performance computing and communications, pp 400–405
3. Ezpeleta J, Colorn JM, Martinez J (1995) A Petri net based deadlock prevention policy for flexible manufacturing systems. IEEE Trans Robot Autom 11(2):303–306
4. Shan Z, Lin C, Ren F et al (2003) Modeling and performance analysis of a multiserver multiqueue system on the grid. In: Proceedings of the ninth IEEE workshop on future trends of distributed computing systems (FTDCS'03), pp 800–804
5. Li Z, Liu B, Li X, Yin Y (2011) Verification of safety-critical software requirement based on Petri-net model checking. Syst Eng Electron 33(2):458–463
6. Yuan C, Huang Y, Zhao W, Li X (2011) A study on fairness of place/transition systems-to make fairness fairer. Trans Inst Meas Control 33(1):50–58
7. Tang Z, Yin J (2012) Extended Petri net-based simulation modeling and analysis. J Univ Electron Sci Technol China 41(1):153–157
8. Zhang M, Duan Z (2012) Synthesis rules of WF-nets and their application in process design. J Xidian Univ (Nat Sci) 239(2):678–782
9. Petri CA (1975) Interpretations of net theory. ISF-Report75–07,GMD, St.Augustin, F.R.G., pp 700–802
10. Zhang J, Wu Z (1995) Computation of synchronic distance in marked T-graph. J SE Univ 33(5):100–103
11. Wu Z (2006) Introduction to Petri nets. China Machine Press, Beijing
12. Wang LL, Wu ZH, Huan F (2008) Computation of synchronic distance in marked T-net. Comput Sci 35(10):100–103
13. Zeng QT, Wu ZH (2002) The language characters analysis of analogous S-graph. Comput Sci 29(5):120–122
14. Duan H, Zeng QT (2004) Analysis on the liveness of S-net. Mini-Micro Syst 25(11):1975–1978

Petri Net Modeling and Analysis Based on Gene Logic Network

Yulin Zhang, Shudong Wang, Hongyue Wu and Yan Yi

Abstract Petri net has recently emerged as a promising tool for the modeling and analysis of molecular networks. In this research, gene logic network constructed by approach-logic analysis of phylogenetic profiles method is described. In order to depict the logic interactions between genes, a new Petri net formalism with augmented inhibitor arc is proposed, which is called ALTPN. Two different types of places and different transitions are formulated in ALTPN, and then corresponding firing rule is given. Further, ALTPN of all 1-order and 2-order gene logic types are listed. Finally reachability graph method is used to accomplish asynchronous dynamic analysis of gene logic interactions in colon cancer and some conclusions are drawn.

Keywords Gene · Logic network · LAPP · Petri net · Reachability graph

1 Introduction

Understanding the regulatory mechanisms of gene expression is one of the key problems in molecular biology. Some kinds of models have been studied to express gene regulatory networks such as Boolean networks [1, 2], differential equations [3, 4], Bayesian networks [5, 6] and Petri nets [7] et al. One method to depict such regulatory networks is the logical approach initially developed by Thomas [8]. In fact, this logical formalism can be regarded as 1-order logic, just depicting logic interaction such as activation or repression between two genes. But this is not enough to describe intracellular interaction "network" complexity, so in order to identify more complex logical relationship, a computational approach-logic analysis

Y. Zhang · S. Wang (✉) · H. Wu · Y. Yi
College of Information Science and Engineering, Shandong University
of Science and Technology, Qingdao, 266510 Shandong, China
e-mail: wangshd2008@yahoo.com.cn

of phylogenetic profiles (abbreviated as LAPP) was proposed, which was applied into 4873 distinct orthologous protein families of 67 fully sequenced organisms [9]. Since then, the theory and application of higher logic have been made some progresses. For example, a three-way gene interaction model that captured the dynamic of co-expression relationships between two genes was described [10]. Then LAPP method was improved and concept of support and confidence in the judgment of logical types was proposed, further the logical network of 16 active genes at Arabidopsis bud in different external stimulus conditions are constructed [11]. Obviously, the network generated using LAPP has higher-order logics and shows the internal structure of the complex biological systems in more details.

Petri net (PN) offers complementary framework to analyze the structure and dynamical properties of biological systems. Indeed, Petri net modeling has already been applied to various types of biological networks, mainly focusing on biochemical pathways, signaling pathways, gene regulation et al. [12–16]. For example, Chaouiya et al. [12] proposed a systematic rewriting of logical genetic regulatory graphs in terms of standard Petri net models, further they analyzed isolated regulatory circuits and confirmed fundamental dynamical properties. Comet et al. [13] developed a translation of multi-valued genetic logical regulatory networks into high-level Petri net, accounting for observed biological behaviors such as homeostasis, multi stationarity, or even more specific temporal properties. Matsuno et al. [15] have considered hybrid Petri net modeling of gene regulatory networks, providing a convenient way to represent protein concentration dynamics being coupled with discrete switches.

The paper is organized as follows. Firstly logical network constructed by the LAPP method is described and the dynamical behavior of logical network is defined. Then a new augmented Petri net formalism with augmented inhibitor arc is proposed, which is called ALTPN. Two types of places and different transitions are formulated in ALTPN. Based on the firing rule, reachability graph method is used to accomplish asynchronous evolution of ALTPN. Finally, a biological example using ALTPN to simulate logic interactions is analyzed.

2 Petri Net Modeling Based on Gene logic Network

2.1 Structure and Dynamic of Gene Logic Network

According to LAPP method based on gene expression profile, gene logic network is constructed and defined as U-logic network in the paper [11]. In fact, U-logic network is a weighted and directed graph which represents 1-order or 2-order logic interactions between genes by uncertainty coefficient. Each logic interaction involves two gene sets: the source set and the target set. Indeed, one gene can be the target of several logic interactions. Figure 1 shows an example of one 1-order logic and one 2-order logic interaction.
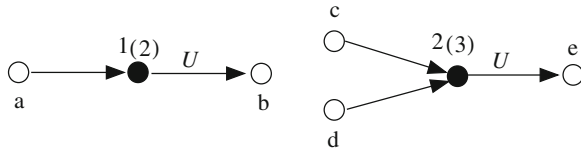
Fig. 1 *White circles* denote genes and while *black circles* denote intermediate nodes. Number 1(2) indicates that gene a regulates gene b by logic type 2 of 1-order and 2(3) indicates gene c and gene d together regulate gene e by logic type 3 of 2-order

In Boolean network, state of nodes is driven by exclusive logic function while the dynamic rule of U-logic gene network is different from that of Boolean network. Assume the network has n gene nodes, Let the state of gene i at moment t be 0 or 1, without loss of generality, the triple set $\left\{ (k, U_{k \to i}, f_k^1) | k \in \{1, 2, \ldots, n\}, k \neq i \right\}$, which denotes that gene k regulates gene i by the logic function f_k^1 and the uncertainty coefficient is $U_{k \to i}$ at moment t; Similarly the triple set $\left\{ \left((k_1, k_2), U_{k_1,k_2 \to i}, f_{k_1,k_2}^2 \right) | k_1, k_2 \in \{1, 2, \ldots, n\}, k_1, k_2 \neq i, k_1 \neq k_2 \right\}$ denotes that gene k_1, k_2 regulate gene i by the proper function f_{k_1,k_2}^2 and the uncertainty coefficient is $U_{k_1,k_2 \to i}$ at moment t; In this research, we consider that the state of gene i at moment $t+1$ is related to the states of adjacent nodes. If the truth-value of proper function $f_{k_1,k_2,\ldots,k_i}^i (i \in \{1, 2\})$ with independent variable being the state of genes $k_1, k_2, \ldots, k_i$ at moment t is 1, then these i genes $k_1, k_2, \ldots, k_i$ activate the expression of gene j. let $\underset{k_1,k_2,\ldots,k_i \to j}{\mathrm{Sign}}(t) = 1$. Otherwise, these i genes $k_1, k_2, \ldots, k_i$ inhibit the expression of gene, Let $\underset{k_1,k_2,\ldots,k_i \to j}{\mathrm{Sign}}(t) = -1$. Then the state of gene j at moment $t+1$ is determined by transition rule

$$
x_j(t+1) = \begin{cases} 1, \sum_{Z_{j1}} U_{k \to j} \underset{k \to j}{\mathrm{Sign}}(t) + \sum_{Z_{2j}} U_{k_1,k_2 \to j} \underset{k_1,k_2 \to j}{\mathrm{Sign}}(t) \geq 0.5 \\ 0, \sum_{Z_{j1}} U_{k \to j} \underset{k \to j}{\mathrm{Sign}}(t) + \sum_{Z_{2j}} U_{k_1,k_2 \to j} \underset{k_1,k_2 \to j}{\mathrm{Sign}}(t) < 0.5. \end{cases} \tag{1}
$$

According to above transition rule, the state of network evolves by synchronous or asynchronous assumptions. Detailed introduction about U-logic network can be seen in [11].

2.2 Petri Net Modeling

For the U-logic network, state of gene node takes value {0, 1} and state transition is driven by (1) instead of being driven by logic function merely in Boolean network. Therefore, it should be interesting to articulate the U-logic network with an augmented PN formalism at the stage.

Firstly we analyze the basic form that a target gene node in the U-logic network is affected by one logic interaction such as 1-order logic function or 2-order logic function with corresponding uncertainty coefficient.

Some requirements and constrains must be given as follows:

(1) Places which are marked by white circles denote gene nodes. Place which is marked by black circle denotes intermediate node and it just performs auxiliary function to get logic truth, so it is called supplementary place.
(2) Transitions represented by white rectangle correspond to logic function. Transitions represented by rectangle with vertical lines correspond to transition rule of gene node in formula (1). So two types of transitions are called logic transition, threshold transition respectively.
(3) Token denotes state of genes. Here consider Boolean case of genes, so the expression state of each gene takes its value in $\{0, 1\}$. The number of tokens in each place should be limited to 1 and this constraint can be satisfied by adding a capacity restriction for these places. For supplementary place, because it has not biological significance, so its initial token is 0.
(4) There is no consumption of the tokens in the places of input genes obviously, it just activate the transcription of target gene, but are not consumed. So this constraint is satisfied by adding read arcs for these places.

All logic types (including 2 1-order logic and 2 2-order logic types) and corresponding Petri net formalism are listed in Table 1. For all logic types, the unified definition of Petri net is given.

Definition 1 Let $N = <P, T, F, K, U, W, M_0>$ is 7-tuple, where $P \cap T = \varnothing$, $P \cup T \neq \varnothing$.

(1) $P = P_1 \cup P_2$ is a set of places. P_1 represents the class of ordinary places. Each ordinary place corresponds to a gene. P_2 represents the class of supplementary places which do not have the actual biological meaning, just perform transition

Table 1 Petri nets formalism corresponding to 4 logic types with uncertainty coefficient

Order	Logic type	Petri net
1	$B = \neg A$	
1	$B = A$	
2	$C = A \wedge B$	
2	$C = A \vee B$	

function to decide the state of next nodes based on the rules of logical network dynamics.

(2) $T = T_1 \cup T_2$ is a set of transitions. T_1 represents the class of logic transitions which are defined logic transitions which control state transformation from source gene places to supplementary places. T_2 represents the class of special transitions which are defined threshold transitions which determine state transformation from supplementary places to target gene.

(3) $F = I \cup F_1 \cup F_2$, I is a set of normal inhibitor arcs and $I \subseteq P \times T$. F_1 is a set of standard arcs and $F_1 \subseteq (P \times T) \cup (T \times P)$. F_2 is a set of augmented inhibitor arcs and only satisfies $F_1 \subseteq (P_2 \times T_2)$.

(4) $K : P \to I$ is a capacity function and for each place $p \in P, K(p) = 1$.

(5) $U : F \to R$ is a weight function of arcs. For each augmented inhibitor arc, the value of the function is same (between 0 and 1). We call it gene U-logic factor. Otherwise, the default values of the function for other types of arcs are 1.

(6) $W : T_2 \to R$ is a function of each threshold transition. The value of the function is between 0 and 1. In the paper, the value equals to 0.5. The default threshold values are 1 for logic transitions.

(7) $M_0 : P \to \{0, 1\}$ is the initial marking.

For the logic transition, $t \in T_1$ is enable if and only if the following condition is established: $\forall p \in {}^{\bullet}t$, if $<p, t>$ is inhibitor arc, then $M(p) = 0$; if $<p, t>$ is standard arc, then $M(p) = 1$. For the threshold transition, $t \in T_2$ is enable if and only if the following condition is established: $\forall p \in {}^{\bullet}t$, if p is supplementary place and $<p, t>$ is a standard arc, then $\sum f(M(p)) \cdot U \geq \Gamma$; if p is supplementary place and $<p, t>$ is an augmented inhibitor arc, then $\sum f(M(p)) \cdot U < \Gamma$; if p is ordinary place, then $M(p) = 1$, where $f = \begin{cases} 1, & M(p) = 1 \\ -1, & M(p) = 0 \end{cases}$. If the logic tran-

sitions and the threshold transition can enable at the same time, the logic transitions must have priority of enabling. Execution rule of transitions is given as follows. If the transition $t \in T_1$ is enabled at M, $M[t > M']$ and

$$M'(p) = \begin{cases} M(p) + 1, & \textit{if } p \in t^{\bullet} \\ M(p) - 1, & \textit{if } p \in {}^{\bullet}t \\ M(p), & \textit{if } <p, t> i \textit{ is an inhibitor arc} \\ M(p), & \textit{others} \end{cases}.$$

Obviously, firing rule is composed of enable rule and execution rule. The augmented PN can be called ALTPN in the paper.

3 Reachability Analysis of Petri Net

Obviously ALTPN actually is threshold formalism with augmented inhibitor arc, so it is more suitable for using reachability graph than incidence matrix analyze its dynamic. Here we present one biological illustrations of rewriting of gene logic

Fig. 2 ALTPN formalism corresponding to two gene logic interactions
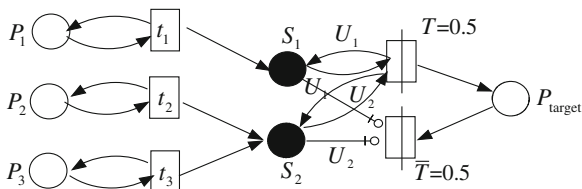
regulatory networks into ALTPN. In colon cancer, the vascular endothelial growth factor VEGFC and cancer gene K-RAS together regulate apoptosis-related cysteine peptidase CASP9. CASP9 is present if and only if VEGFC or K-RAS is present and expression of gene APC may cause up-regulation of CASP9 [17]. Assume uncertainty coefficient of 1-order logic function is $U_1 = 0.4$ and uncertainty coefficient of 2-order logic function "or" is $U_2 = 0.838$, then corresponding Petri net is given in Fig. 2.

CASP9 is the target gene of two logic interactions. One source gene set includes VEGFC and K-RAS, the other set includes APC. In Fig. 2, P_1, P_2 and P_3 denote gene APC, VEGFC, K-RAS respectively. P_{target} denotes target gene CASP9. Transition t_1 denotes 1-order logic function "=". Transition t_2 and t_3 denote 2-order logic function "or" respectively. S_1 and S_2 denote supplementary place. Here suppose states of gene nodes evolve on the basis of formula (1) by asynchronous assumption.

In Fig. 3, 1 or 0 marking with horizontal line is the state of supplementary places. If $M_0 = (1, 1, 0, 0)$, according to firing rule that if the logic transition and the threshold transition can enable at the same time, the logic transition must have priority of enabling. There are two firing sequences of transition: $t_1 \rightarrow t_2 \rightarrow T$ or $t_2 \rightarrow t_1 \rightarrow T$, then final state is $M = (1, 1, 0, 1)$. If $M_0 = (1, 1, 1, 1)$, t_1, t_2, and t_3 threshold transition $\overline{T}$ can fire at same time, but in the work, transition t_1, t_2, and t_3
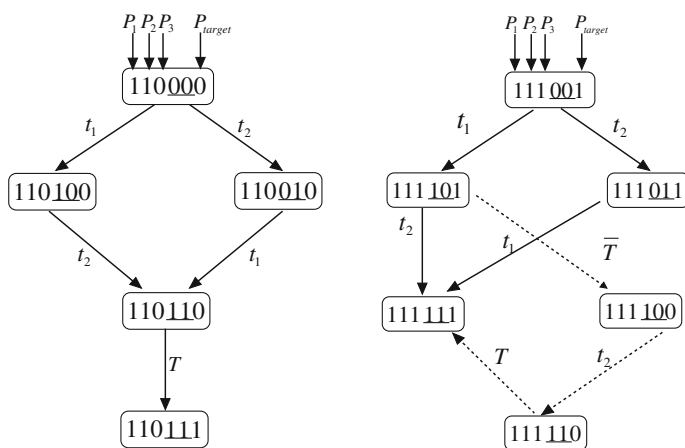
Fig. 3 Reachability graph from two different initial state to final state

has priority of firing. After t_1, t_2, and t_3 fire, the token of each supplementary place changes to 1, then $\overline{T}$ and T can not fire again, so the final state is $M = (1, 1, 1, 0)$. In Fig. 3, $t_1 \rightarrow \overline{T} \rightarrow t_2 \rightarrow T$ with dotted arrow actually does not exist. Finally, we can draw some conclusions. (1) There may exist more than one path from initial state evolve to final state in reachability graph, but there must be a shortest path conforming to the firing rules we give. (2) No transitions can fire again so the final state is a deadlock state for the example.

4 Conclusion

In this work, we use PN depict higher order logic interactions between genes, proposing ALTPN formalism with different types of places, transitions and augmented inhibitor arcs. In fact, ALTPN can be synthesized by logic and threshold Petri net. Then we analyze one biological illustrations of rewriting of gene logic regulatory network into ALTPN and draw some conclusions.

References

1. Akutsu T, Miyano S, Kuhara S (1999) Identification of genetic networks from a small number of gene expression patterns under the Boolean network model. Pac Symp Biocomputing 4:17–28
2. Kauffman SA (1972) The large-scale structure and dynamics of gene control circuits: an ensemble approach. Theory Biol 44(1):167–190
3. Marlovits G, Tyson CJ, Novak B, Tyson JJ (1998) Modeling M-phase control in Xenopus oocyte extracts: the surveillance mechanism for unreplicated DNA. Biophys Chem 72:169–184
4. Novak B, Csikasz-Nagy A, Gyory B, Chen K, Tyson JJ (1998) Mathematical model of the fission yeast cell cycle with checkpoint controls at the G1/S, G2/M and metaphase/anaphase transitions. Biophys Chem 72:185–200
5. Friedman N, Linial M, Nachman I (2000) Using Bayesian networks to analyze expression data. J Comput Biol 7:601–620
6. Grzegorczyk M, Husmeier D, Rahnenfuhrer J (2011) Modeling non-stationary gene regulatory processes with the BGM model. Comput Stat 26(2):199–218
7. Chaouiya C (2007) Petri net modeling of biological networks. Briefings Bioinf 8(4):210–219
8. Thomas R (1973) Boolean formalization of genetic control circuits. J Theor Biol 42(3):563–585
9. Bowers PM, Cokus SJ, Eisenberg D, Yeates TO (2004) Use of logic relationships to decipher protein network organization. Science 306(5706):2246–2249
10. Zhang J, Ji Y, Zhang L (2007) Extracting three-way gene interactions from microarray data. Bioinformatics 23(21):2903–2909

11. Wang S, Chen Y, Wang Q, Li E, Su Y, Meng D (2010) Analysis for gene networks based on logic relationships. J Syst Sci Complexity 23(5):435–447
12. Chaouiya C, Remy E, Ruet P, Thieffry D (2004) Qualitative modelling of genetic networks: from logical regulatory graphs to standard petri nets. Lect Notes Comp Sci 3099:137–156
13. Comet J, Klaudel H, Liauzu S (2005) Modeling multi-valued genetic regulatory networks using high-level petri nets. Lect Notes Comp Sci 3536:208–227
14. Hardy S, Pierre NR (2004) Modeling and simulation of molecular biology systems using petri nets: modeling goals of various approaches. J Bioinf Comput Biol 2(4):619–637
15. Matsuno H, Doi A, Nagasaki M (2000) Hybrid petri net representation of gene regulatory network. Pac Symp Biocomputing 5:338–349
16. Simao E, Remy E, Thieffry E, Chaouiya C (2005) Qualitative modeling of regulated metabolic pathways: application to the tryptophan biosynthesis in E. coli. Bioinformatics 21(2):190–196
17. Ruan X, Wang J, Li H (2007) The establishment and application of a logical network model based on gene expression profiles of genes. Prog Biochem Biophys 34(8):871–880

DOA Estimation for Nonuniform Linear Arrays Using Root-MUSIC with Sparse Recovery Method

Xinpeng Du, Xiang Xu and Lizhi Cheng

Abstract Direction-of-arrival (DOA) estimation with nonuniform linear arrays (NLA) using the sparse data model is considered. Different with the usually used sparse data model, we introduce a linear interpolation operator which can transform the data of the NLA to the data of a virtual uniform linear array (VULA). We first reduce the dimension of the model using the singular value decomposition technique, next recover the solution of the reduced MMV using a compressed sensing (CS) algorithm, then get the data of the VULA using the recovery result and the linear interpolation operator, and lastly use root-MUSIC to estimating DOA. The method is called CS-RMUSIC. The experiments illustrate the good efficiency of the CS-RMUSIC algorithm.

Keywords Direction-of-arrival · Non-uniform linear array · Compressed sensing · Root-MUSIC · Array interpolation

1 Introduction

Direction-of-arrival (DOA) estimation is an important part in sensor array signal processing, and is widely used in many applications including radar, sonar, radio astronomy, mobile communication, and so on [1, 2].

The array manifold of the uniform linear array (ULA) is the Vandermonde matrix, and has very good properties, and then ULA is an important form of sensor arrays in the literature of DOA estimation. However, in practical cases, the sensors cannot be uniformly spaced in line due to the limitations of the array platform and the consideration of the cost. A good remedy for these problems is constructing the

X. Du (✉) · X. Xu · L. Cheng
Department of Mathematics and System Science, College of Science, National University of
Defense Technology, Changsha 410073 Hunan, People's Republic of China
e-mail: duxp1984@yahoo.com.cn

nonuniform linear arrays (NLAs) [3]. Luckily, the performances of NLAs are much better than ULAs with the same number of sensors [1]. Most of traditional algorithms for DOA estimation are based on ULAs, and they cannot be directly applied to NLAs. Although MUSIC (MUltiple Signal Classification) and root-MUSIC can still be applied to NLAs [3], but their performances both are poor.

Array interpolation (AI) [4–6] is a classical and extensively used technique for estimating DOA with NLA. The AT technique obtains the data of a virtual ULA (VULA) by imposing a linear interpolation operator on the data of a real NLA, with interpolator coefficients selected to minimize the interpolation error for a signal arriving from a given angular sector [6]. After obtaining the VULA data, all classical DOA estimation algorithms, such as root-MUSIC and ESPRIT (Estimating Signal Parameter via Rotational Invariance) [1, 2], can be applied. And we call the whole algorithms as AI-RMUSIC [5] and AI-ESPRIT [4], respectively. It has been proved that AI is a very effective technique, but it cannot identify the sources outside the interpolation sector [1] (see Fig. 1).

With the developing of the CS theory [7, 8], the DOA estimation problem can be cast as a sparse signal representation problem [9, 10]. These algorithms are the high resolution algorithms and can be applied both into ULAs and NLAs. However, the sparse data model is established using an angular grid, and these algorithms are limited by the angular grid.

In this paper, based on the sparse data model [9] and the AI technique, we propose a new algorithm to estimate the DOAs with NLA. Similar to the AI technique, but different from CS methods, the proposed algorithm can obtain the data of a VULA (corresponding to the NLA). Additionally, similar to the CS
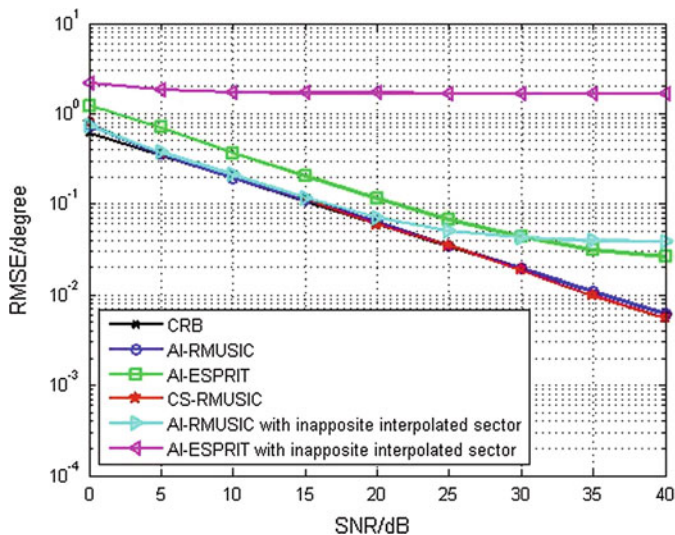
Fig. 1 DOA estimation RMSEs versus SNR for $\theta = 10°$, $J = 100$, and $D_{LA} = \begin{bmatrix} 0 & 1 & 3 \end{bmatrix}\frac{\lambda}{2}$

methods, but different from the AI technique, the method can also estimate the waveform of the sources. The algorithm firstly recovers the solution of the sparse model, then obtains the VULA data using a linear interpolation operator, and lastly applies any classical DOA estimation algorithm (e.g. root-MUSIC) into the VULA data to estimate the DOAs. The experiment results show that the proposed algorithm is efficient and can overcome the limitation of the AI technique.

The outline of the rest for the paper is as follows. In Sect. 2, a modified sparse data model for problem (1) is proposed. The new algorithm based on the CS method and the AI technique is presented in Sect. 3. In Sect. 4, the numerical experiments are implemented to evaluate the performance of the proposed algorithm. Lastly, we draw some conclusions in Sect. 5.

2 The Data Model

In this section, we propose a modified sparse data model for DOA estimation with NLA. Before introducing the proposed model, we first introduce the real data model and the sparse data model.

2.1 The Real Data Model

Throughout the paper, we consider a linear array (either ULA or NULA), labelled as $D_{LA} = [d_1, \ldots, d_M] \frac{\lambda}{2}$, of M sensors $(M < \lceil d_M \rceil \triangleq m)$, where $d_M > \cdots > d_1 = 0$ and they may not be integers, $\lceil \cdot \rceil$ denotes the ceiling operation, λ denotes the minimum wavelength of the received signals. We assume that L far-field narrowband sources impinging on the NLA come from the distinct directions $\theta = [\theta_1, \ldots, \theta_L]^T$. Then the received signal at the t-th measurement of the received signals takes the form

$$y(t) = As(t) + n(t) \quad t = 1, \ldots, J \tag{1}$$

where the vector $y(t) \in \mathbb{C}^{M \times 1}$ represents the received signal, the matrix $A = [a(\theta_1), \ldots, a(\theta_L)] \in \mathbb{C}^{M \times L}$ is the array manifold, the vector $a(\theta_i) = [1, \ldots, e^{-j\pi d_M \sin \theta_i}]^T \in \mathbb{C}^{M \times 1} (1 \leq i \leq L)$ is the steering vector of θ_i, the vector $s(t) = [s_1(t), \ldots, s_L(t)]^T \in \mathbb{C}^{L \times 1}$ is the sources vector, and $n(t) \in \mathbb{C}^{M \times 1}$ denotes a complex additive white Gaussian noise vector. The superscript 'T' denotes transpose.

To express the model with matrix form, let $Y = [y(1), \ldots, y(T)] \in \mathbb{C}^{M \times J}$, $S = [s_1, \ldots, s_L]^T \in \mathbb{C}^{L \times J}$, $s_i = [s_i(1); \ldots; s_i(T)] \in \mathbb{C}^{1 \times J} (1 \leq i \leq L)$ and $N = [n(1), \ldots, n(T)] \in \mathbb{C}^{M \times J}$ is the noise matrix. Then the matrix form of (1) can be written as

$$Y = AS + N. \tag{2}$$

2.2 The Sparse Data Model

Now, we introduce sparse data model [9, 10] of problem (2). In the beginning of estimation, we don't know the true values of $\{\theta_l\}_{l=1}^L$. However, we can always determine the angular sector(s) (in the array's view field) which includes $\{\theta_l\}_{l=1}^L$, let Ω denote the angular sector(s). Suppose $\Delta = \{\varphi_k\}_{k=1}^K$ be a grid that covers Ω, where K is the number of potential DOAs and $K \gg L$. Let the notation $a(\varphi)$ denote the steering vector of the sensor array corresponding to the direction φ, then we define $A_\Delta^r = [a(\varphi_1), \ldots, a(\varphi_K)]$. We assume the grid is so fine that each $\theta_l(l = 1, \ldots, L)$ is equal or very close to some $\{\varphi_j\} j = 1, \ldots, K$. Then the sparse data model of the DOA estimation can be expressed as:

$$Y = A_\Delta^r S_\Delta + \bar{N} \tag{3}$$

where $S_\Delta = [s_1, \ldots, s_K]^T$ are the unknown sampling signals imping on the array and $\bar{N}$ is the sum of N and the modeling error. The matrix S_Δ is jointly sparse, that is, only a small number of rows of S_Δ are nonzero. If there exists $\theta_{l'}$ on the grid such that $\varphi_{k'}$ is equal to $\theta_{l'}$ or closer to $\theta_{l'}$ than others, the k'-th row of S_Δ should be nonzero. Then the DOA estimation problem is transformed to finding the nonzero rows of S_Δ.

2.3 The Proposed Model: Sparse Data Model with VULA

As we know, many algorithms have been proposed to estimate DOAs with ULA. Even if some algorithms, such as MUSIC and root-MUSIC, can be applied to non-uniform arrays, but the results are not very good. In order to make use of the good property of ULA, some techniques have been developed to transform the data of NLA to the data of VULA, followed by the classical method of DOA estimation. Array interpolation (AI) is one of these techniques and is extensively used. Motivated by the AI technique, we propose a new technique to solve the problem (1). Firstly, we define the VULA of D_{LA} as a ULA with m sensors spaced $\frac{\lambda}{2}$ apart. Then we suppose a matrix G satisfy $A_\Delta^r = G A_\Delta^v$, where A_Δ^v is the array manifold of the VULA. The same as the AI technique, the matrix G can be defined as

$$G = A_\Delta^r (A_\Delta^v)^H \left\{ A_\Delta^v (A_\Delta^v)^H \right\}^{-1}. \tag{4}$$

Generally, the matrix G is of full row rank, and rank $(G) = M$. If the real sensor array is a ULA, G is an identical matrix. Invoking above equations into Eq. (3), we can obtain the new sparse data model

$$Y = G A_\Delta^v S_\Delta + \bar{N}. \tag{5}$$

3 The Proposed Method

Problem (5) is an MMV (Multiple Measurement Vectors) [9, 11, 12]. In the CS theory, l_1-SVD [9], M-OMP (Orthogonal Matching Pursuit for MMV) [11] and compressive MUSIC [12] etc. have been proposed to solve the MMV problem. The complexity is large when J is large. To reduce the complexity, we use the SVD (singular value decomposition) of Y. SVD can decompose a matrix into the signal subspace and the noise subspace, and then one can keep the signal subspace and recast the problem in a lower dimension. This technique is firstly introduced in [9] in the context of MMV, and it also can reduce the sensitive to noise. Mathematically, the idea can be expressed as follows.

When $J \leq L$, we consider problem (5) directly. When $J > L$, suppose the SVD of Y is $Y = ULV^H$, where the superscript 'H' denotes conjugate transpose. Keep a reduced $M \times L$ dimensional matrix Y_{sv} which contains most of the signal power $Y^{sv} = ULD_k = YVD_k$, where $D_k = [I_k \ 0]$. Here, I_k is an $L \times L$ identity matrix. Let $S_\Delta^{sv} = SVD_k$ and $\bar{N}^{sv} = \bar{N}VD_k$. Then problem (4) can be recast as

$$Y^{sv} = GA_\Delta^v S_\Delta^{sv} + \bar{N}^{sv}. \tag{6}$$

Then we use the MMV algorithms to solve problem (6), such as M-OMP and compressive MUSIC. And we can obtain the estimator of the support of S_Δ or S_Δ^{sv} (e.g. using M-OMP), denoted by $\tilde{I}$; or the estimator of the waveforms $\tilde{S}_\Delta$ or $\tilde{S}_\Delta^{sv}$ (e.g. using l_1-SVD). We can use the recovery results of MMV algorithms to estimate the DOAs and its waveforms directly. This method is very efficient, but it is limited by the angular grid [9].

In this paper, in order to use the good property of ULA and overcome the limitation of the sparse algorithms for DOA estimation, we use the recovery result to estimate the VULA data. That is, we obtain the virtual array data by $\tilde{Y} = A_\Delta^v \tilde{S}_\Delta$ or $\tilde{Y} = A_\Delta^v \tilde{S}_\Delta^{sv}$. Then we apply the classical algorithms into the data $\tilde{Y}$ to estimate the DOAs. In this paper, we apply the root-MUSIC algorithm into the data $\tilde{Y}$ to estimate DOAs, and we call the proposed algorithm as CS-RMUSIC.

4 Experiments

In this section, we give two examples to study the performances of CS-RMUSIC algorithm. We assume that the angular sector $[\theta_l, \theta_r]$ of interest is known and is uniformly divided by $0.01°$ in the AI technique, and the angular sector $\left[\frac{-\pi}{2}, \frac{\pi}{2}\right]$ is divided into $2N + 1$ angles uniformly in the sparse data model. The number of independent trials is set to be $n = 500$. The noises added to the signals are all assumed to be the complex additive white Gaussian noises. The ESPRIT algorithm in AI-ESPRIT is the least square ESPRIT. The performances of CS-RMUSIC are compared with the Cramer-Rao Bound (CRB), AI-RMUSIC and AI-ESPRIT.

The CRB [13] on the variance of estimation error of θ is defined as

$$CRB(\theta) = \frac{\sigma^2}{2} \left\{ \sum_{t=1}^{J} \mathcal{R}\left[X^H(t)D^H\left[I - A(A^H A)^{-1} A^H \right] D X^H(t) \right] \right\}^{-1} \quad (7)$$

where $D = [d(\theta_1), \ldots, d(\theta_L)]$, $d(\theta) = \frac{\partial a(\theta)}{\partial \theta}$, $X(t) = diag[s_1(t), \ldots, s_L(t)]$, and $\mathcal{R}[\cdot]$ denotes the real part of the expression in the square brackets.

The Root Mean Square Error (RMSE) for the estimator of $\theta_i (1 \leq i \leq L)$ is defined by the quantity $\mathrm{RMSE}(\theta_i) \triangleq \sqrt{\frac{1}{n_i} \sum_{k=1}^{n_i} \left| \hat{\theta}_{ik} - \theta_i \right|}$, where n_i is the number of non-aberrant estimator for θ_i and $\hat{\theta}_{ik}$ is the k-th $(1 \leq k \leq n_i)$ estimator of θ_i.

4.1 Example 1

Suppose $M = 3, D_{LA} = \begin{bmatrix} 0 & 1 & 3 \end{bmatrix} \frac{\lambda}{2}$ and a narrowband far-field single source comes from $10°$ In the AI technique and CS-RMUSIC, the VULA has four sensors with array aperture $m = 3$. In the sparse data model, the angular space $\left[\frac{-\pi}{2}, \frac{\pi}{2} \right]$ is equally divided by $1°$ (i.e. $N = 90$). Suppose $[\theta_l, \theta_r] = [0°, 20°]$ in the AI technique. The number of snapshots is set to be $J = 100$.

In Fig. 1, we display the DOA estimation RMSEs of the CS-RMUSIC, AI-RMUSIC and AI-ESPRIT algorithms versus the SNR ranging from 0 to 40 dB, compared with the CRB. Figure 1 also illustrates the results of AI-RMUSIC and AI-ESPRIT when the interpolated sectors are inappositely selected at $[20°, 40°]$.

From Fig. 1, CS-RMUSIC is an efficient algorithm. It is shown that CS-RMUSIC performs much better than AI-ESPRIT, and that CS-RMUSIC performs a little better than AI-RMUSIC at high SNR. Because CS-RMUSIC does not need to determine the interpolated sectors, the poor estimations resulted from inapposite selected interpolated sectors will not appear. This means that CS-RMUSIC can overcome the limitation of the AI technique.

4.2 Example 2

Suppose an NLA $D_{LA} = \begin{bmatrix} 0 & 1 & 2 & 6 & 7 & 10 & 13 \end{bmatrix} \frac{\lambda}{2}$ of 7 sensors, and two narrowband uncorrelated far-field sources come from the directions of $10°$ and $20°$. The number of snapshots is set to be $J = 200$. In the AI technique and the CS-RMUSIC algorithm, the VULA has fourteen sensors with array aperture $m = 13$. In CS-RMUSIC, the angular space $\left[\frac{-\pi}{2}, \frac{\pi}{2} \right]$ is equally divided by $0.25°$ (i.e. $N = 360$). Suppose $[\theta_l, \theta_r] = [0°, 30°]$ in the AI technique. The DOAs estimation RMSEs of CS-RMUSIC, AI-RMUSIC and AI-ESPRIT versus SNR ranging from 0 to 40 dB for $10°$ are shown in Fig. 2, compared with the CRB.
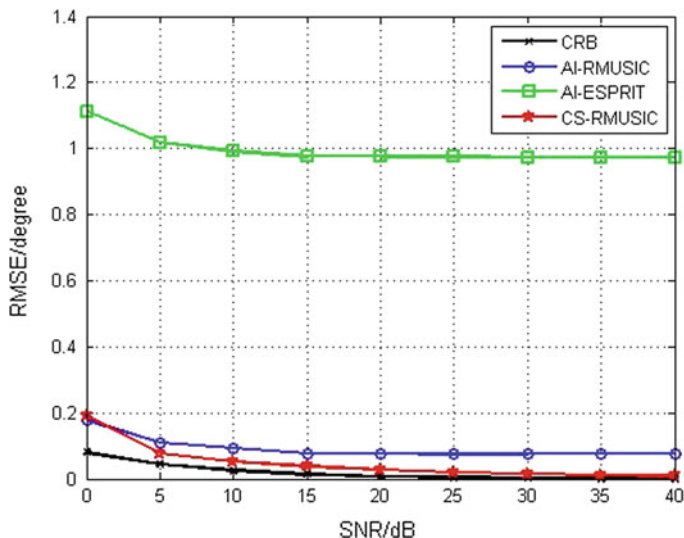
Fig. 2 DOA estimation RMSEs versus SNR for $D_{LA} = \begin{bmatrix} 0 & 1 & 2 & 6 & 7 & 10 & 13 \end{bmatrix} \frac{\lambda}{2}$, $\theta = [10°, 20°]$, $J = 200$

Figure 2 shows that the performances of CS-RMUSIC for estimating the two sources are much better than the performances of AI-ESPRIT and better than the performances of AI-RMUSIC.

From the above examples, we find that the proposed CS-RMUSIC algorithm is efficient, that it performs better than AI-ESPRIT and AI-RMUSIC, and that it can overcome the limitations of the AI technique.

5 Conclusion

Based on the developing CS theory and the classical AI technique, we propose the CS-RMUSIC algorithm for DOA estimation with NLA in this paper. The CS-RMUSIC algorithm first obtains the recovery solution of the sparse data model, and then uses linear interpolation operator and the recovery solution to obtain the VULA data, lastly applies the root-MUSIC algorithm into the VULA data to estimate the DOAs. The numerical experiments show that the CS-RMUSIC algorithm is efficient, and it performs better than AI-ESPRIT and AI-RMUSIC. Moreover, the CS-RMUSIC algorithm can overcome the limitations of the AI technique.

References

1. Tuncer TE, Friedlander B (2009) Classical and modern direction-of-arrival estimation. Academic Press, New York
2. Van Trees HL (2002) Optimum array processing. Wiley, NewYork
3. El Kassis C, Picheral J, Mokbel C (2010) Advantages of nonuniform arrays using root-MUSIC. Signal Process 90:689–695
4. Weiss AJ, Garish M (1991) Direction finding using ESPRIT with interpolated arrays. IEEE Trans Signal Process 39:1473–1478
5. Friedlander B (1993) The root-MUSIC algorithm for direction finding with interpolated arrays. Signal Process 30:15–25
6. Friedlander B, Weiss AJ (1992) Direction finding using spatial smoothing with interpolated arrays. IEEE Trans Aerosp Electron Syst 28:574–587
7. Donoho DL (2006) Compressed sensing. IEEE Trans Inf Theory 52:1289–1306
8. Candès EJ, Wakin MB (2008) An introduction to compressive sampling. IEEE Signal Process Mag 25:21–30
9. Malioutov D, Cetin M, Willsky AS (2005) A sparse signal reconstruction perspective for source localization with sensor arrays. IEEE Trans Signal Process 53:3010–3022
10. Liu ZM, Huang ZT, Zhou YY, Liu J (2012) Direction-of-arrival estimation if non-circular signals via sparse representation. IEEE Trans Aerosp Electron Syst 48:2690–2698
11. Chen J, Huo X (2006) Theoretical results on sparse representations of multiple measurement vectors. IEEE Trans Signal Process 54:4634–4643
12. Kim JM, Lee OK, Ye JC (2012) Compressive MUSIC: revisiting the link between compressive sensing and array signal processing. IEEE Trans Inf Theory 58:278–301
13. Stoica P, Nehorai A (1989) MUSIC, maximum likelihood and Cramer-Rao bound. IEEE Trans Acoust Speech Signal Process 37:720–741

Research of ROM Based on Molecular Beacon DNA Computing

You-rui Huang, Jing Wang and Xiao-min Tian

Abstract ROM is an indispensable part in the DNA computer system. It is used to store binary data. Based on DNA computing, the ROM developed by molecular beacon can realize the data storage in a high speed and on a large scale. And it has direct influence on the development process of DNA computer technology. This paper puts forward a new implementation method of logic gates, which based on the molecular beacon calculation mode, and realize a simple combinational logic circuit, the construction of the ROM. In this method, logic gates are represented by molecular beacons. Input signals are represented by single strands of DNA, which can realize the operations of logic gates in DNA type; the connection between two logic gates according to the circuit levers can be realized by the marks of tubes. The construction of the ROM, which use the molecular beacon and based on DNA computing, will play a huge role in advancing the development of DNA computer.

1 Introduction

The development of traditional integrated circuits has become increasingly close to the limits of technology. The electronic circuits based on semiconductor technology will be difficult to meet the density and complexity requirements of

---

Y. Huang (✉) · J. Wang · X. Tian
Anhui University of Science and Technology, Huainan, China
e-mail: hyr628@163.com

J. Wang
e-mail: jingwang319@163.com

X. Tian
e-mail: tianxiaomin100@163.com

computer circuits. Therefore, constructing the DNA computer which has great parallelism and high-speed computing power will be of great significance for the implementation of artificial intelligence [1]. The research in the field of DNA computing has attracted researchers' attention in the field of mathematics, computer science, life science. Molecular computing is the key to realizing DNA computing [2]. Since hairpin structure and specific performance of molecular beacon can solve problems effectively [3], and it is easy to operate, and the sensitivity is quite high, it can be used as an effective carrier of DNA computing.

In the electronic computer design, microprocessor using basic logic gates to build electronic circuits. Similarly, in molecular computing, one of the main goals is to design molecular logic gates [4, 5], which can be addressed. Molecular switch and logic gates have obvious similarities. In various types of molecular switches, fluorescence has switched stands out since its high sensitivity, simple operation, easy to test the output signal characteristics. ROM is a permanent memory that can only read the data stored in advance. It is commonly used in the storage of a variety of fixed programs and data. And it has the advantage of not easy loss and high reliability. If humans can use biological means to construct DNA molecular ROM, then it is possible to solve the problem that is storing data with high-speed, large-scale, low-power. And it has an important significance to the development of DNA computers.

This paper combines the advanced molecular beacon chip technology and proposes a model which used the molecular beacon to realize ROM. The proposition of this model makes it possible to solve the data-storage problem by using molecular beacon to calculate. By designing the sequence structure of molecular beacon skillfully and controlling the state open or closed of the structure of hairpin effectively, the logic gates are simulated with molecular beacons. Use single strands of DNA as input, through melting of the double chain and processing the calculated results in accordance with the conditions, make the output level directly as the input to the next level. For the integration of the logic circuit, using the method of classification mark to structure a 2-4 decoder. By using methods such as marking the molecular beacons with different color fluorescence [6], PCR and hybridization, the storage arrays and output circuit of ROM are simulated. Enable signal is controlled by heating the solution to dissolve the double-stranded DNA.

2 Molecular Beacons

2.1 The Structure and Working Principle of the Molecular Beacons

The molecular beacon is a stem-loop double-labeled probe with oligonucleotides, which is a hairpin structure. The structure of molecular beacon [6] is shown in Fig. 1. It consists of three parts, the loop for identification, the stem, and the

Fig. 1 The structure of molecular beacon
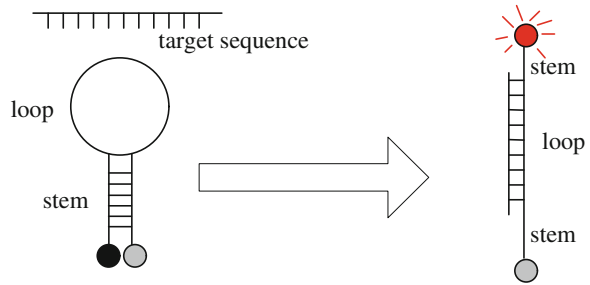
fluorescent molecule and quenches molecule which are connected to both ends of the stem sequence. When it is in free-state, the fluorescent molecule and quench molecule are very near. The fluorescence emitted and distributed in the form of heat when the fluorescent molecule absorbed by quench molecule. Fluorescence is quenched almost complete. Therefore, it can be detected. When using certain methods, the identical sequence of the loop can hybridize with the target sequence specially, and the double-stranded hybrids will form. The fluorescent molecule and quench molecule increase distance, making the fluorescence detected [7, 8].

2.2 The Constructor Methods of Logic Gates Using Molecular Beacons

In this paper, the logic gates working process are divided into two processes, calculation and output. In the calculation process, logic gates calculate the input signals. The output process is realized by adding enzyme and some biological operations.

2.2.1 The Encoding Principle of Logic Gates

Add the molecular beacons present logic gates into tubes, and it consists of two parts: the sequences X1 and X2, which represents the input signal, and the stem sequences S and S$'$ on its two sides. S and S$'$ will form a hairpin structure in appropriate conditions. The two input signals of the logic gate are respectively represented by a DNA sequence of length 2l. For each input signal X, if X $=$ 1, then add X, else if X $=$ 0, then don't add X. The loop sequence of the molecular beacon is represented by the complement chain links of X1 and X2. The requirements to encode the stem sequence S and S$'$ are: ① for the AND gate, the binding force of the stem must be greater than that of any 2l−length sequence of the loop, while lesser than that of 4l−length sequence of the loop when they are hybridizing; ② for the OR gate and NON gate, the binding force of the stem must be lesser than that of any 2l−length sequence of the loop when they are hybridizing.

Fig. 2 Four input conditions of the AND gate
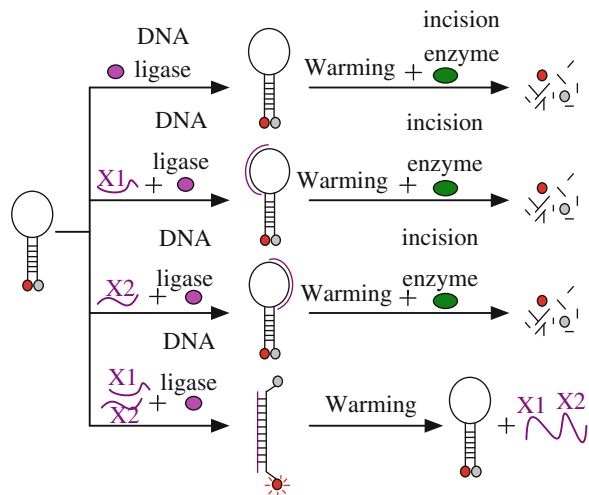
2.2.2 Biological Realizable Methods of Specific Logic Gates

The AND gate is shown in Fig. 2. According to the encoding principle, only when X1 = 1 and X2 = 1 and the DNA connection enzyme exists (attention: DNA connection enzyme only works when X1 = 1 and X2 = 1), did the binding force of the loop sequence greater than that of the stem sequence. The molecular beacon would open, and the red fluorescence would be emitted. If the red fluorescence is detected, heat the solution to unlock the double-stranded DNA, then elute and purify the solution, the same single-stranded DNA with the input signal can be gotten as an input to the next level of logic gates (i.e., X1 = 1 and X2 = 1). Otherwise, heat to unlock the double-stranded DNA and add incision enzyme corresponding with X1 and X2, to make X1 and X2 destroyed, then elute and purify the single-stranded DNA will not be gotten, i.e., The output is X1 = 0 and X2 = 0. Take the two cases: ① X1 = 1 and X2 = 0; ② X1 = 1 and X2 = 1, for example, the concrete operational process is shown in Fig. 3.

The NON gate is shown in Fig. 4. According to the encoding principle, only when both of input X1 and X2 are 0, did the binding force of the loop sequence lesser than that of the stem sequence. The molecular beacon would not open, and the blue fluorescence would not be emitted. If the blue fluorescence is detected, heat the solution to unlock the double-stranded DNA and add incision enzyme corresponding with X1 and X2, to make X1 and X2 destroyed, then elute and purify, the single-stranded DNA would not be gotten as input to the next level of logic gates (i.e., X1 = 0 and X2 = 0). Otherwise, add X1 or X2, after the blue fluorescence is detected, heat the solution, then elute and purify, then the single-stranded DNA would be gotten as an input to the next level of logic gates (i.e., X1 = 1 or X2 = 1). Take two cases: ① X1 = 1 and X2 = 0; ② X1 = 0 and X2 = 0, for example, the concrete operational process is shown in Fig. 5.

Fig. 3 Biological realizable operation of the AND gate
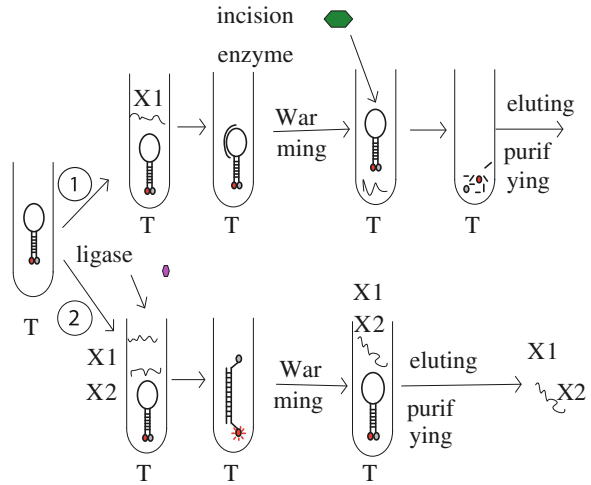
Fig. 4 Four input conditions of the NON gate
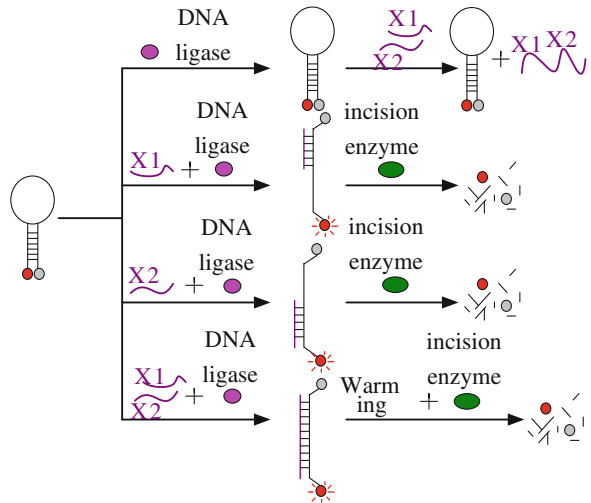
3 The Structure and Working Principle of ROM

In general, the memory is composed of three parts, the storage arrays, the address decoder and the output control circuit. The storage array is composed of a number of storage units, each memory cell storing a binary data. Usually storage units are arranged in matrix form, and are organized into groups according to a certain number of bits. Each time read a set of data, that is, a 'word'. In order to distinguish different words, give a number to each word, called the address. The address decoder decodes the address code into the corresponding control signal of word units. Then the control signal selected the specific storage unit from the

Fig. 5 Biological realizable operation of NON gate
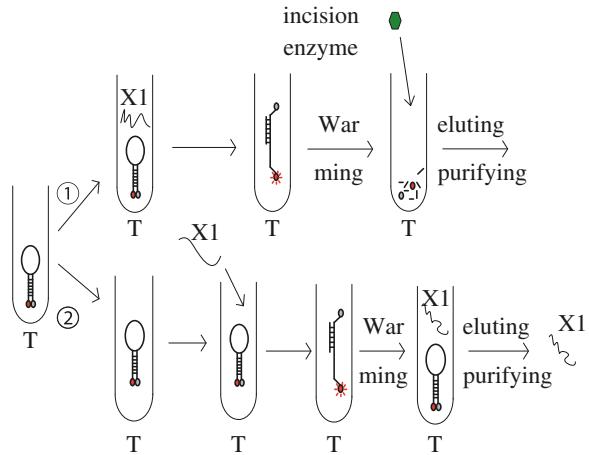
storage matrix, and the data in the storage matrix are sent to the output control circuit. Take the ROM, which has 4 'word' and the length of each word is 4, for example. Its specific structure is shown in Fig. 6.

It can be seen from the diagram above. When the enable signal is active, for the given address X1, X2, only one output of the decoder is 0; the corresponding word line is low. The diodes cross at the word line and all bit lines conduct, make the corresponding bit line output go low. Then the results can be gotten from the output buffer.

Fig. 6 The schematic diagram of ROM structure
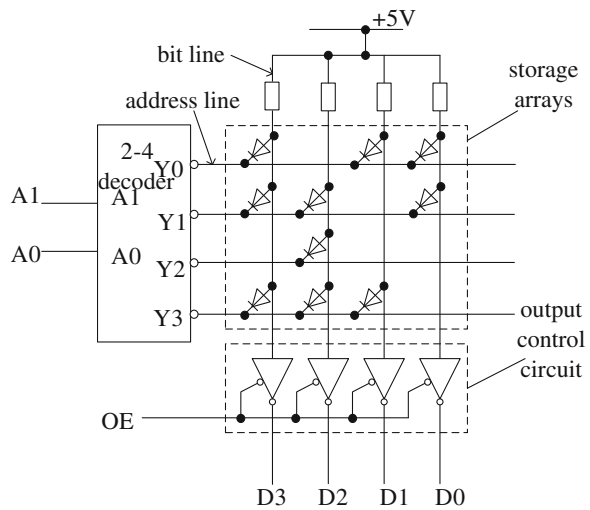
4 The Implementation Methods of ROM

In order to construct the ROM using molecular beacons, this paper realizes the function of each part, and finally completes the function of accessing and reading data. For convenience, regulate that when the decoder output high level, it is effective. That is, when output single-stranded DNA, the corresponding address is selected. By detecting the fluorescence, which is emitted after the hybridization between the output of single-stranded DNA and the molecular beacon, determine the selected address. The output of the single-stranded DNA is amplified by PCR, and then adding them into the test tubes which represent the corresponding address. Simulate the storage arrays and the output circuit by the hybridization between the single-stranded DNA and molecular beacons in the test tubes. These test tubes which represent the same address (word line) are all added fluorescent molecules marked by different colors. The first principle of adding is to make the rule that these molecular beacons represent diodes on different bit lines respectively, according to their different colors. And then determine whether to add the corresponding molecular beacons according to whether there is a diode at the cross of each word line and bit line. Thus, only the tubes on behalf of the selected address output single-stranded DNA. And the single-stranded DNA hybridizes with molecular beacons and then emits fluorescence, then the output can be gotten. In order to realize the function of the enable signal, control the output through whether heating the solution to make the target sequence and molecular beacons melt or not.

4.1 Algorithm Steps to Construct the Decoder

In DNA-based bio-computer, how to realize the connection between the logic gates is a relatively difficult problem. This paper uses the single-stranded DNA to represent the input and output signal. Use the method that marked the tube according to the circuit level to realize the connection of molecular logic gates without affecting the work of molecular beacons. It improves the accuracy and the reliability of operation greatly.

Take the 2-4 decoder, for example. In order to simplify, here ignores the enable signal. That is, it has been in working condition. The logic circuit is shown in Fig. 7.

Mark the tube in order to solve the problem of calculating according to the steps. Add the output of each layer into the test tube of the next level. For example, a logic circuit with the layer of k. Pour the output of the test tube T_{i-1} into a test tube T_i which in the layer i ($1 < i < k$). Finally, get the output of the k layer as a result. Each layer of a logic circuit may contain several logic gates. Each of them can be denoted as T_i^j ($j = 1, 2 \ldots n$). Specific steps are as follows:

1. Add the molecular beacons encoded for logic gates into test tubes, the codes of the three logic gates are shown in Table 1. (Black is for the loop part and purple is for the stemencode the stem sequence part).
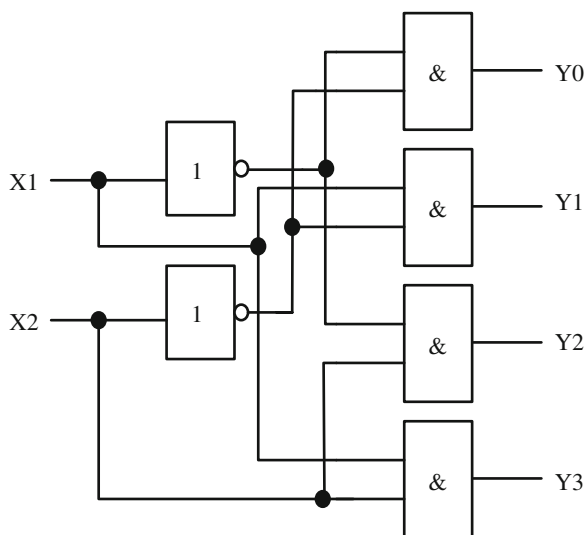
Fig. 7 The logic circuit of the 2-4 decoder

Table 1 The codes of the three logic gates

Logic gate	The DNA code of molecular beacon	Fluorescent material (color)
AND	5'-ATTCGTACTTCAGCTTGCATGAAT-GTACGAAT-3'	Dylight594 (red)
OR	5'-CCGTTTAATTCAGCTTGCATGAAT-TTAAACGG-3'	Dylight488 (green)
NOT	5'-GGCATTATTTCAGCTTGCATGAAT-ATAATGCC-3'	Dylight405 (blue)

2. Select the types of logic gates in accordance with the equivalent circuit. And mark the test tubes based on the circuit level. There are three layers in the 2-4 decoder, mark them respectively as T_1^1, T_1^2, T_2^1, T_2^2, T_2^3, T_2^4, T_3^1, T_3^2, T_3^3, T_3^4. The 10 test tubes are added with the molecular beacons to represent two NON gates, four AND gates, and four NON gates in proper order (Fig. 8).

3. Encode the input signal with the sequence complementary to the loop sequence of the molecular beacon, i.e., AAGTCGAACGTACTTA. When the input is 1, enter the sequence; while the input is 0, do not add it;

4. Add the input X1 to T_1^1 and X2 to T_1^2, then output according to the method of the NOT gate;

5. Pour the outputs of tubes T_1^1 and T_1^2 into T_2^1, add the input X1 and the output of the tube T_1^2 into T_2^2, add the input X2 and the output of the tube T_1^1 into T_2^3, then add the input X1 and X2 to T_2^4;

6. Elute and purify the outputs of T_2^1, T_2^2, T_2^3, T_2^4. And observe whether each of them contains a single strand of DNA. If there is, then record 1; else, record 0. And they are recorded as Y_0, Y_1, Y_2, Y_3 in turn. Note that $Y_i = 1$ (i = 0, 1, 2, 3) indicates an effective output signal.
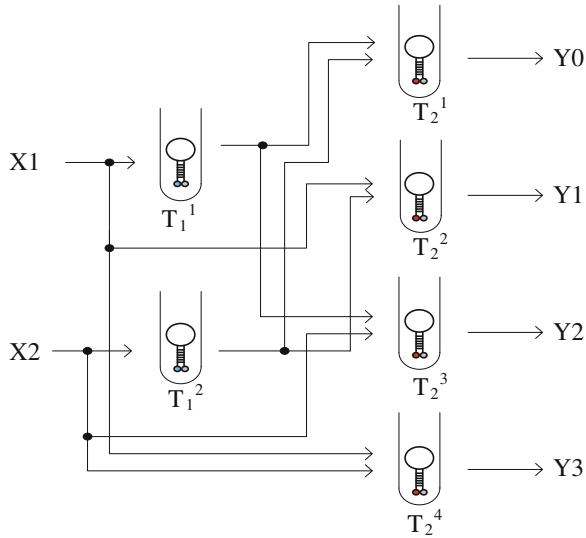
Fig. 8 The realizable circuit of the 2-4 decoder by molecular beacon

Table 2 The experimental results

Input	Does the tube T_2^1, T_2^2, T_2^3, T_2^4 contain single-stranded DNA	Y_0, Y_1, Y_2, Y_3
X1 = 0, X2 = 0	YES, NO, NO, NO	1, 0, 0, 0
X1 = 0, X2 = 1	NO, YES, NO, NO	0, 1, 0, 0
X1 = 1, X2 = 0	NO, NO, YES, NO	0, 0, 1, 0
X1 = 1, X2 = 1	NO, NO, NO, YES	0, 0, 0, 1

The experimental results:

Record the results under various input conditions in Table 2. It can be seen that the decoder realized its decoding function.

4.2 The Implementation of the Storage Array and the Output Control Circuit

It can be seen from the above analysis about the working principles of ROM. When an address is selected, the storage array and the output circuit realize the function accessing and reading the binary data stored in the address commonly. For convenience, it can be understood as long as there is a diode at the cross of the selected word line and bit line, the digital output is 1; otherwise it is 0. This paper use molecular beacons marked with different colors to represent diodes on different bit lines. Add appropriate molecular beacons into corresponding test tubes according to whether there are diodes exist in the chart or not, which is

equivalent to writing data to the storage. The hybridization between the output single-stranded DNA of decoder and molecular beacon makes the hairpin structure turn on, and the fluorescence would be emitted. Then the data can be read according to the color of the fluorescence.

Specific steps are as follows:

1. Firstly, four kinds of molecular beacons were labeled with different fluorescent colors, with blue representing the diode D3 on the highest bit, yellow representing the diode D2 on the second-highest bit, green representing the diode D1 on the second-lowest bit, red representing the diode D0 on the lowest bit;

2. The test tubes T_3^1, T_3^2, T_3^3, T_3^4 is corresponding to the decoder output tubes T_2^1, T_2^2, T_2^3, T_2^4, respectively. Add the molecular beacons marked with different colors respectively into the test tubes T_3^1, T_3^2, T_3^3, T_3^4, according to whether the diodes exist at the cross in the chart. In the above chart, add molecular beacons marked with blue, green and red colors to T_3^1, add molecular beacons marked with blue, yellow and red colors to T_3^2, add molecular beacon only marked with yellow colors to T_3^3, add molecular beacons marked with blue, yellow and green colors to T_3^4;

3. The outputs of the tubes T_2^1, T_2^2, T_2^3, T_2^4 are amplified by PCR, and then add the solution to the test tubes T_3^1, T_3^2, T_3^3, T_3^4, and observe the fluorescent colors of the tubes. If the fluorescence emitted suggests that the corresponding bit is 1, otherwise, it is 0.

Take the given address X1X2 = 00 for example, after the selection by the decoder, the address Y0 is selected. Only the tube T21 output the single-stranded DNA. And it hybridizes with the molecular beacons in the tube T31 after it is amplified by PCR. Its biological implementation process is shown in Fig. 9.

Fig. 9 The biological implementation process of storage arrays and the output
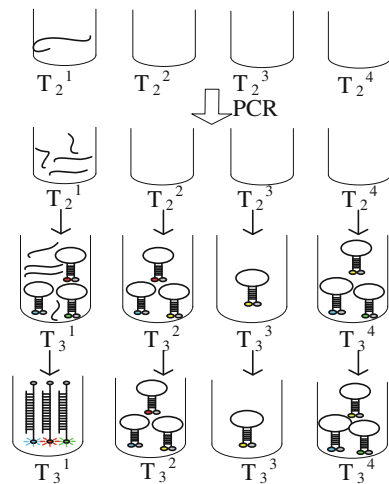
Table 3 The fluorescence solution and output data results

The selected address	The color of fluorescence	The output data $(D_3D_2D_1D_0)$
Y_0	Blue, green, red	1011
Y_1	Blue, yellow, red	1101
Y_2	Yellow	0100
Y_3	Blue, yellow, green	1110

4.3 The Simulation and Experimental Results of the Enable Signal

The role of the enable signal is to read data normally when it is at the low level, and all the outputs are pulled down when it is at the high level. In this paper, simulate the enable signal's work by using the method that heating the solution to make the target sequence and the molecular beacon melt. Specifically, when the enable signal is invalid, heat the solution, the molecular beacon and the target sequence will be melted. And the hairpin structure will restore, the fluorescence will be quenched, then all the outputs are 0; while when the temperature of the solution is reduced, the enable signal is valid, the target sequence and the molecular beacon hybridize, fluorescence is emitted, and the data are output normally. The results: Table 3 recorded the fluorescence solution and the output data they presented when four different addresses are selected. It can be seen that ROM realizes the function to access and read the stored data.

5 Conclusions

This paper proposes a method which uses the molecular beacon to analog logic gates, and constructs a ROM. The advantages are as follows: (1) The design of the DNA sequence of the molecular beacon is skillful, thus make it possible to control the opening and closing of the hairpin effectively. (2) Represent input and output signals with single-stranded DNA, and process the output of every logic gate, then make the output into the next level as an input signal. The operation is simple and reliable. (3) Mark tubes according to the level of Boolean circuits, realize the decoding function, the operations are very convenient. (4) Use fluorescent molecules in different colors instead of diodes, simulate working principle of the storage arrays using the method that the fluorescence is emitted after the PCR and hybridization reaction. This method achieves the function that accessing and reading the stored data simply and reliably, and it provides a guideline for the construction of the large-capacity ROM.

References

1. Qian L, Winfree E (2011) Scaling up digital circuit computation with DNA strand displacement cascades. Science 332:1196–1201
2. Zhang DY, Seelig G (2011) DNA computing and molecular programming. Lect Notes Comput Sci vol 6518:176–186 Springer
3. Adleman LM (1994) Molecular computation of solutions to combinatorial problems. J Sci 266:1021–1023
4. Bi S, Yan YM, Hao SY, Zhang SS (2010) Colorimetric logic gates based on supramolecular DNAzyme structures. J Angew Chem 49:4438–4442
5. Stojanovic MN, Semova S, Kolpashchikov D, Macdonald J, Morgan C, Stefanovic D (2005) Deoxyribozyme-based ligase logic gates and their initial circuits. J Am Chem Soc 127:6914–6915
6. Zhixiang Y, Zhang F, Xu J (2003) DNA computing based on molecular beacon. J Biomath 18:497–502
7. Ogihara M, Ray A (1998) DNA-based self-propagating algorithm for solving bounded-fan-in Boolean circuits. In: Third conference on genetic programming, Morgan Kaufman Publishers, San Francisco, 725–730 (1998)
8. Amos M, Gibbons A, Hodgson D (1999) Error-resistant implementation of DNA computations: DNA based computers II, 44th edn. American Mathematical Society, Providence, pp 151–162

Improved Genetic Algorithm for Solving Optimal Communication Spanning Tree Problem

Nguyen Duy Hiep and Huynh Thi Thanh Binh

Abstract Optimal Communication Spanning Tree (OCST) is a well-known NP-hard problem on the graph that seeks for the spanning tree with the lowest cost. The tree cost depends on the communication volume between each pair of nodes. This paper proposed an improved Genetic Algorithm combining with Ahujia and Murty's Tree Improvement Procedure. The proposed algorithm was experimented on known benchmark tests which used in many papers related to OCST problem, and random instances from 200 to 500 vertexes. The experimental results show that the proposed algorithm is better than the heuristic and out-performance the most recent evolutionary algorithm approaches.

Keywords Optimal communication spanning tree · Genetic algorithm

1 Introduction

The Optimal Communication Spanning Tree (OCST) problem (also known as Minimum Communication Spanning Tree Problem or Simple Network Design Problem) was introduced by Hu in [1]. It arises in many applications such as design of wire-based communication networks under quality of service requirements, transportation network, and VLSI design problem.

The OCST problem can formally be defined as follows. Given a weighted, undirected graph $G(V, E)$ with the number of nodes $n = |V|$, and the number of

N. D. Hiep (✉) · H. T. T. Binh
School of Information and Communication Technology, Hanoi University of Science and Technology, Hanoi, Vietnam
e-mail: hiepnd@soict.hut.edu.vn

H. T. T. Binh
e-mail: binhht@soict.hut.edu.vn

edges is $|E|$. The distance and the communication demand between each pair of nodes is given by matrix $D = d_{ij}$ and $R = r_{ij}$, respectively. A tree $T(V, E_T)$ where $|E_T| = |V| - 1$, and $E_T \subseteq E$ is a spanning tree on G if it connects all the nodes. The communication cost between two nodes is calculated as multiplication of the communications requirement to the distance between the two nodes in T. The cost $c(T)$ is the total communication cost summed over all pairs of nodes.

$$c(T) = \sum_{i,j \in V} r_{ij} \times d\left(p_{i,j}^T\right)$$

where $d\left(p_{i,j}^T\right)$ is the length of the unique path from node i to node j in T. T is minimum communication spanning tree if $c(T) < c(T')$ for all other spanning trees T'. If no communication requirements r_{ij} are considered then OCST problem becomes Minimum Spanning Tree (MST) problem.

The OCST problem is NP-hard [2]. Furthermore, Reshef showed that the problem is MAX SNP-hard, a class of optimization problem in which there is no approximation scheme unless $P = NP$ in [3].

In our approach, we combine a genetic algorithm with tree improvement procedure introduced by Ahuja and Murty in [4]. Experimental results on known benchmark and random instances show that our algorithm out-performance previous evolutionary algorithms and performance more stable and better than known heuristic methods, especially on large instances.

The rest of this paper is organized as following. In the next section related works is discussed. Section 3, the detail of our improved genetic algorithm is presented. In Sect. 4, we implement our experimental on test instances and compare our algorithm with some known best approaches. The paper concludes with Sect. 5 with some discussions on the future extension of this work.

2 Related Works

The OCST problem was first introduced by Hu in [1], two exact algorithms were also proposed for solving two special case of the problem: the Optimum requirement spanning tree $(d = 1)$ can be solved by using the Ford Fulkeson labeling procedure in $O(n^4)$ time, and the Optimum distance spanning tree $(r = 1)$ also known as minimum routing cost spanning tree MRCT.

Ahuja and Murty introduced both exact and heuristic algorithm for solving OCST problem in [4]. The exact algorithm based on the branch and bound, the lower bound is computed in $O(n^4)$ time and the upper bound is found by the heuristic algorithm. The heuristic algorithm is a two phase algorithm: tree building and tree improvement algorithm.

Palmer and Kerhembaum showed that the good solution for OCST often have a star topology and they presented a heuristic algorithm to exploits it. The algorithm cost $O(n^2)$ to evaluate all-star topology trees present a heuristic. A MST is created

if the star search process is not successful. They also proposed a genetic algorithm in [5] as an alternative approach. Chromosome in Genetic Algorithm (GA) is encoded by a method called Link and Node bias (LNB) that use a vector of bias values for both edge and nodes. The result found by GA is better than by the heuristics algorithm.

Recent approaches for solving OCST problem focus on Evolutionary Algorithms. Sang Moon introduced a GA which uses a sequence length $2(n-1)$ of nodes to represent a spanning tree, called the Cycle-Breaking Tree Construction Routine (CB-TCR) in [6]. The new GA with CB-TCR encoding method shows great improvement, and outperforms other GAs using previous encodings when experiment on the known benchmark instances.

A memetic algorithm, combined of a genetic algorithm with a local search, modified version of Ahuja and Murty heuristic algorithm, was first introduced by Fischer in [7]. In this modified algorithm, limited node sets with different criteria such as: node demand, edge distance, number of edges to node, path lengths, node index… The new memetic algorithm performed better than previous evolutionary algorithms, and running fast on large instances.

Another approach using PSO [8], hybrid PSO [9], these PSO algorithms are better than previous GA, but their performance decelerate quickly on large instances. This paper proposes a new approach for solving OCST—combine a regular genetic algorithm with tree improvement procedure introduced in [4]. Comparing with the best known algorithms, the results show that our algorithm performance better, and much more stable. Especially, our proposed algorithm breaks scores on the Berry35u instance.

3 Improved Genetic Algorithm for Solving Optimal Communication Spanning Tree Problem

Genetic Algorithms are nature-inspired algorithms that try to improve an initial population of solutions by iteratively (in terms of generations) applying a set of operators to the population. It has proven effective on many practical NP-hard problems. Much works research on NP-hard problem, particularly in problems relating to tree has been done.

In this section, an Improved Genetic Algorithm (IGA) with new crossover and mutation operators based on the idea of tree improvement procedure in [4] is proposed. The brief of tree improvement procedure is introduced in the beginning of this section; the remaining is the detail of our algorithm.

3.1 Tree Improvement Procedure

The tree improvement in Ahuja and Murty [4] is a local search heuristic performing 1-exchange move. It takes a spanning tree T as an input, examines each

tree edge e on T by trying to cut it (the tree is cut into two components S and $\bar{S}$), and replace it by candidates edges from $(S \times \bar{S})$. If there is an edge e' on $(S \times \bar{S})$ (one vertex on S and another on $\bar{S}$) makes the cost of T smaller, it will be update on current tree T (e' will replace for e on T), otherwise the procedure will stop.

For examining edge quickly, the two parameters w_i and h_i are calculated. Where w_i is the sum of demands from node i to all nodes in other component, and h_i is component's external traffic's cost if all traffic from i's component to other component routed via node i.

$$w_i = \begin{cases} \sum_{j \in \bar{S}} r_{i,j} & if\ i \in S \\ \sum_{j \in S} r_{i,j} & if\ i \in \bar{S} \end{cases} \quad h_i = \begin{cases} \sum_{j \in S} w_j \times d(p_{i,j}^s) & if\ i \in S \\ \sum_{j \in \bar{S}} w_j \times d\binom{\bar{s}}{i,j} & if\ i \in \bar{S} \end{cases}$$

3.2 Improved Genetic Algorithm

Individual representation: Each individual is a spanning tree, represented by a list of edges.

Crossover operator: The crossover operator takes two spanning trees $T_1(V, E_1)$ and $T_2(V, E_2)$ as input. A new graph $G'(V, E')$ where $E' = E_1 \cup E_2$ is created and then apply tree improvement procedure on T_1 and T_2 using edges on G' to create two children T_1' and T_2'.

Mutation operator: In the mutation operator, we use the same strategy as in crossover operator. From the input tree $T(V, E_T)$, we create a new graph $G'(V, E')$ using all edges in E_T and randomly selected edges in $\{E \backslash E_T\}$. Then the tree improvement procedure is applied on T to create T' (Only edges on G' are considered).

The Improved Genetic Algorithm (Fig 1).

Random spanning tree is used for creating initial population because it can prevent fast convergence. In each generation, the number of time crossover and mutation operations performed are based on the crossover rate Pc and mutation rate Pm.

Fig. 1 Improved genetic algorithm

The Improved Genetic Algorithm

```
Procedure IGA
Begin
P ← Init population
      While ¬ terminate do
      Begin
            For i=1 to m
                  P'←crossover(P)
            End for
            For  j = 1 to k
                  P''←mutation(P')
            End for
      P←P''
            End
End procedure
```

The termination condition of the algorithm is one of the following:

- The number of generations reaches the fix number,
- The best value of population is not improved after some fixed generations.

4 Experimental Results

4.1 Problem Instances

We use two data tests for testing the proposed algorithm; the first is the known benchmark instances with number of nodes range from 6 to 100, which were used in many papers related to OCST problem [3, 6–10]. The second data test is Non-Euclidean graphs, randomly constructed, with the number of nodes are 200, 300 and 500.

4.2 Experiment Setup

We experiment our proposed algorithm independently and compare its performance to Ahuja and Murty heuristic (AM-H) in [4], HGA using Node Bias (NB) and CB-TCR encoding methods in [9], and two GA variants (GA1 and GA2) presented by Raidl and Justrom and experimented by Rothlauf in [11]. For GA1, we use the GA1(MST) results, which initialize population by minimum spanning tree.

The parameters used for our algorithm in the experimental are: the number of generation is 20, population size is 5, crossover rate Pc is 80 %, mutation rate Pm is 20 %. Each instance is run 20 times and average values are used for discussion.

Our algorithm installed in Java language, run on a machine with Intel Core i5, RAM 3 GB, Windows 7 Professional.

4.3 Experimental Results

Table 1 shows the results of our algorithm run on two data sets: the known benchmark and the random test. The experiment results in Table 1 shows that on all the known benchmark instances, our algorithm can find the best known values, and the standard deviation value is 0, or lower than 4 % of best known value in all instances. The lower value of standard deviation shows that our algorithm is stable. The algorithm runs quite fast on small and acceptable on large instances. It takes only 22 s for instance Raild100, 13 min for random200, and 45 min for random300 instances.

Table 1 The results found by IGA on all the problem instances over 20 runs

Instance	Best known	IGA			
		Best	Diff (%)	Avg	Std
Berry6	534	**534**	0	**534**	**0**
Berry35	16915	**16915**	0	**16915**	**0**
Berry35u	16167	**16167**	0	16169.7	7.2
Palmer6	693180	**693180**	0	**693180**	**0**
Palmer12	3428509	**3428509**	0	**328509**	**0**
Palmer24	1086656	**1086656**	0	**1086656**	**0**
Raidl10	53674	**53674**	0	53674	**0**
Raidl20	157570	**157570**	0	157570	**0**
Raidl50	806864	**806864**	0	806864	**0**
Raidl75	1717491	**1717491**	0	1717491	**0**
Raidl100	2561543	**2561543**	0	2562686.9	4783.4
Random200	–	**23540014250**	–	23540014250	**0**
Random300	–	20929836119	–	20930181783	**1545857**
Random500	–	**39254783470**	–	39256292590	**1891179**

– means the result is not available

Best the best result found after 20 running times. *Diff* (%) = [(best—best known) * 100 %] / best known. *Avg* the average of results calculated on 20 running times

Std the standard deviation of the results calculated on 20 running times

Table 2 shows the comparison with other algorithms. On all known benchmark instances, the IGA algorithm gives better result than all current best evolutionary algorithms for solving OCST problem at present. Especially, on the Berry35 u instance, it can find the best known value while the others cannot.

Table 2 Comparison the best results found by IGA and GA1, GA2 [11], PSO [8] and HGA [9] on the known benchmark instances over 20 runs

Instance	Best known	GA1 (MST)	GA2	HGA (NB)	HGA (CB-TCR)	PSO	IGA
Bery6	534	534	534	534	534	534	**534**
Berry35	16915	16915	16195	16915	16915	–	**16915**
Berry35u	16167	–	–	19704	16329	–	**16167**
Palmer6	693180	693680	693180	693180	683180	693180	**693180**
Palmer12	3428509	3720657	3623253	3428509	3428509	3428509	**3428509**
Palmer24	1086656	1154449	1250073	1086656	1086656	1138360	**1086656**
Raidl10	53674	57141	55761	53674	53674	53674	**53674**
Raidl20	157570	159911	158974	157570	157570	157570	**157570**
Raidl50	806864	852091	880927	806864	806864	826499	**806864**
Raidl75	1717491	1971638	2003433	1717491	1719849	–	**1717491**
Raidl100	2561543	2831167	2935381	2561543	2580665	**–**	**2561543**

GA1, GA2 two GA variants presented by Raidl and Justrom and experimented by Rothlauf in [11]. *HGA* (*NB*) Hybrid GAs in [9] using node bias encoding, *HGA* (*CB-TCR*) Hybrid GAs in [9] using Cycle-Breaking Tree Construction Routine (CB-TCR) encoding

Table 3 Comparison the results between IGA and AM-H [4] over 20 runs

Instance	Best know	Cost		Diff (%)		Avg		Std	
		AM-H	IGA	AM-H	IGA	AM-H	IGA	AM-H	IGA
Berry6	534	534	534	0	0	534	534	0	0
Berry35	16915	16915	16915	0	0	16915	16915	0	0
Berry35u	16167	16225	**16167**	0.36	0	20016.65	**16169.7**	1044.07	**7.2**
Palmer6	693180	693180	693180	0	0	693180	693180	0	0
Palmer12	3428509	3541915	**3428509**	3.3	0	3549967.2	**3428509**	1895.29	**0**
Palmer24	1086656	1086656	1086656	0	0	1087564.2	**1086656**	213.76	**0**
Raidl10	53674	53674	53674	0	0	53674	53674	0	0
Raidl20	157570	157570	157570	0	0	157642	**157570**	24.62	**0**
Raidl50	806864	806864	806864	0	0	806864	806864	0	0
Raidl75	1717491	1717491	1717491	0	0	1721262	**1717491**	1651.56	**0**
Raidl100	2561543	2736957	**2561543**	6.8	0	2747966.15	**2562686.9**	**4767.91**	4783.4
Random200	–	23613787475	**23540014250**	–	–	23948896208	**23540014250**	737744218.7	**0**
Random300	–	20929836119	20929836119	–	–	20963202726	20930181783	65271109.41	**1545857**
Random500	–	39295546827	**39254783470**	–	–	40997475726	**39256292590**	945660706.6	**1891179**

Table 3 shows the comparison the results between IGA and AM-H [4]. The experiment results show that our algorithm have better performance than AM-H on all test instances.

5 Conclusion

In this paper, we introduce a novel approach to solve the OCST problem. We conduct experiment to evaluate our proposed algorithm and compare its performance to HGA using CB-TCR, NB encoding methods [9], AM-H [4] and GA1, GA2 [11]. The experimental result shows that our proposed algorithm outperforms all of above algorithm on almost instances. In future work, we are planning to experiment the algorithm with different initial constructions, and parallelize it to improve the performance on large instances.

References

1. Hu TC (1974) Optimum communication spanning trees. SIAM J Comput 3:188–195
2. Johnson DS, Lenstra JK, Kan AHGR (1978) The complexity of the network design problem. Networks 8:279–285 Winter 1978
3. Reshef E (1999) Approximating minimum communication cost spanning trees and related problems. In Master's thesis, Feinberg Graduate School of the Weizmann Institute of Science, Rehovot 76100, Israel
4. Ahuja RK, Murty VVS (1987) Exact and heuristic algorithms for the optimum communication spanning tree problem. Transp Sci 21(3):163–170
5. Palmer CC, Kershenbaum A (1994) Representing trees in genetic algorithms. In: Proceedings of the 1st IEEE conference on evolutionary computation, vol 1, IEE Service Center, Piscataway, NJ, pp 379–384
6. Soak SM (2006) New evolutionary approach for the optimal communication spanning tree problem. IEICE Trans 89(10):2882–2893
7. Fischer T, Merz P (2007) A memetic algorithm for the optimum communication spanning tree problem. In: Bartz-Beielstein T, Aguilera M, Blum C, Naujoks B, Roli A, Rudolph G, Sampels M (eds) HM: 4th international workshop on hybrid metaheuristics, vol 4771. Springer, Berlin, pp 170–184
8. Hoang AT, Le VT, Nguyen NG (2010) A novel particle swarm optimization-based algorithm for the optimal communication spanning tree problem. In: Proceedings of the 2010 2nd international conference on communication software and networks, pp 232–236, Feb 2010
9. Kien PT, Hiep ND, Binh HTT (2011) New hybrid genetic algorithm for solving optimal communication spanning tree problem. In: The 26th symposium on applied computing, Taiwan, pp 1076–1081

10. Rothlauf F (2006) Representations for genetic and evolutionary algorithms, 2 edn, Springer
11. Rothlauf F (2009) On optimal solutions for the optimal communication spanning tree problem. Oper Res 57(2):4

Algorithm for Generating Decision Tree Based on Adjoint Positive Region

Jing Gao

Abstract In this paper, all of the three relationships of attribute selection standard based on positive region, based on rough bound and based on attribute dependency are firstly analyzed. At the same time, it is proved that the three kinds of selection attribute standards are equivalent to each other. Furthermore the advantages and disadvantages of algorithm for generating decision tree based on positive region are analyzed. Meanwhile, aiming at these disadvantages, a new selection attribute standard based on adjoint positive region is proposed. The decision tree generated with the new standard of attribute selection has the following characteristics: fewer leaf nodes, fewer levels of average depth, better generalization of leaf nodes. Finally an example is used to illustrate the advantages of this new selection attribute standard.

Keywords Decision tree · Positive region · Adjoint positive region

1 Introduction

Recently, more and more scholars make wide and deep research on ID3 algorithm, which is proposed by professor Quinlan (in Ref. [1]). In fact, ID3 algorithm adopts information plus as attribute selection standard to select all level node's attribute in decision tree. This results in that the most classification information of tested sample can be obtained easily, when every non-leaf node is being tested. Then the sample set is divided into sub set with this kind of attribute. Here the advantage is appears: the entropy's value in the system is smallest, the mean path from this non-leaf node to its successive node is shortest, the mean depth of generated decision

J. Gao (✉)
Information College, Capital University of Economics and Business, Beijing, China
e-mail: gaojinggaoj@126.com

tree is less, the speed and accuracy of classification is promoted. After all, ID3 algorithm has its own scarcity. In order to eliminate ID3's scarcity, many promoted algorithms are proposed in many references. And ID3 algorithm becomes more and more consummate.

In 1982, rough set model was proposed by professor Pawlak, a Polish researcher (in Ref. [2]). Then some scholars introduce rough set theory into decision tree algorithm and propose generated decision tree algorithm based on rough see (in Ref. [3–10]).

In this paper, all of the three relationships of attribute selection standard based on positive region, based on rough bound and based on attribute dependency degree are firstly analyzed. At the same time, it is proved that the three kinds of selection attribute standards are equivalent to each other. Furthermore, by taking attribute selection standard of positive region as delegate, the advantages and disadvantages of algorithm for generating decision tree based on positive region are analyzed. In order to get rid of these disadvantages, a new standard in selecting attributes is proposed in this paper. This standard is based on adjoint positive region. It is assured that the decision tree has few leaf nodes and its leaf nodes have high generalization ability.

2 Relative Analysis

Definition 1 [3] Information system (also called decision tree) is defined as $S = (U, C, D, V, f)$, here $U = \{x_1, x_2, \ldots, x_n\}$ is universe of discourse. And C is condition attribute set, D is decision attribute set, $f : U \times (C \cup D) \to V$ is information function. $F = C \cup D$, $V = \cup V_a, a \in F, V_a$ is universe of a.

Definition 2 [4] In Information system, marked with $S = (U, C, D, V, f), \forall A \subseteq C$, here $U/A = \{A_1, A_2, \ldots, A_m\}$ stands for what U is divided by A.

$\forall X \subseteq U, A_*(X) = \cup\{A_j|A_j \subseteq X\}$, here $A_*(X)$ stands for the lower approximation set that X acts on A in U. And here $A^*(X) = \cup\{A_j|A_j \cap X \neq \varnothing\}$ stands for the upper approximation set that X acts on A in $U. NEG_A(X) = \cup\{A_j|A_j \cap X = \varnothing\}$, here $NEG_A(X)$ stands for the negative region that X acts on A in $U. BND_A(X) = A^*(X) - A_*(X)$, here $BND_A(X)$ stands for the boundary region that X acts on A in U.

Definition 3 [4] In Information system, marked with $S = (U, C, D, V, f)$, $\forall A \subseteq C$, here $U/A = \{A_1, A_2, \ldots, A_m\}$ stands for what U is divided by A. Here, $U/D = \{D_1, D_2, \ldots, D_h\}$ stands for what U is divided by D. Here $BND_A(D) = \bigcup\limits_{D_i \in U/D} (A^*(D_i) - A_*(D_i))$ stands for rough boundary that condition attribute set, marked by A, acts on decision attribute set, marked by D, in U. Here $BND_A(D)$ is briefly called rough boundary.

Definition 4 [7–10] In Information system, marked with $S = (U, C, D, V, f)$, $\forall A \subseteq C$, here $U/A = \{A_1, A_2, \ldots, A_m\}$ stands for what U is divided by A. Here,

$U/D = \{D_1, D_2, \ldots, D_h\}$ stands for what U is divided by D. Here $POS_A(D) = \underset{D_i \in U/D}{\cup} A_*(D_i)$ stands for positive region that condition attribute set, marked by A, acts on decision attribute set, marked by D, in U. Here $POS_A(D) = \underset{D_i \in U/D}{\cup} A_*(D_i)$ is briefly called positive region. And here $\gamma_A(D) = |POS_A(D)|/|U|$ stands for dependence degree that decision attribute set, marked by D, acts on condition attribute set, marked by A, in U. And here $\gamma_A(D)$ is briefly called dependence degree.

Lemma 1 *Suppose that A and B stands for random set respectively, then* $A - B = A \cap \bar{B}$; $\overline{A \cap B} = \bar{A} \cup \bar{B}$; $\bar{\bar{A}} = A$; $\overline{A \cup B} = \bar{A} \cap \bar{B}$; *Here* $\bar{A}$ *stands for complement set of A.*

Theorem 1 *Suppose that* A_1, A_2, B_1, B_2 *stands for random set respectively, and on the condition of* $A_1 \cap B_2 = \varnothing, A_2 \cap B_1 = \varnothing$, *then* $(A_1 - B_1) \cup (A_2 - B_2) = (A_1 \cup A_2) - (B_1 \cup B_2)$. *The following is probative approach*:

$$
\begin{aligned}
(A_1 - B_1) \cup (A_2 - B_2) &= (A_1 \cap \overline{B_1}) \cup (A_2 \cap \overline{B_2}) \\
&= (A_1 \cup A_2) \cap \overline{B_2} \cap \overline{B_1} \cap (\overline{B_1} \cup \overline{B_2}) \\
&= (A_1 \cup A_2) \cap (\overline{B_2 \cup B_1}) \cap (\overline{B_1 \cap B_2}) \\
&= (A_12) \cap (\overline{(B_2 \cup B_1) \cup (B_1 \cap B_2)}) \\
&= (A_1 \cup A_2) \cap (\overline{B_2 \cup B_1}) \\
&= (A_1 \cup A_2) - (B_2 \cup B_1)
\end{aligned}
$$

Theorem 2 *Suppose that* $A_1, \ldots, A_n, B_1, \ldots, B_n$ *stands for random set respectively, and on the condition of* $A_i \cap B_j = \varnothing (i \neq j, i, j = 1, 2, \ldots, n)$, *then* $\overset{n}{\underset{i=1}{\cup}} (A_i - B_i) = \overset{n}{\underset{i=1}{\cup}} A_i - \overset{n}{\underset{i=1}{\cup}} B_i$. *Refer to Theorem 1, above equation can be easily proved.*

Theorem 3 *In Information system, marked with* $S = (U, C, D, V, f)$, $\forall A \subseteq C$, *here* $U/A = \{A_1, A_2, \ldots, A_m\}$ *stands for what U is divided by A.*

And here, $U/D = \{D_1, D_2, \ldots, D_h\}$ *stands for what U is divided by D. Then here is an equation, marked with* $A^*(D_i) \cap A_*(D_j) = \varnothing (i \neq j, i, j = 1, 2, \ldots, h)$. *The following is probative approach* (*reduction to absurdity*):

$$A_*(D_j) = \cup (X | X \in U/A \wedge X \subseteq D_j), A^*(D_i) = \cup (X | X \in U/A \wedge X \cap D_i \neq \varnothing)$$

Suppose $A^*(D_i) \cap A_*(D_j) \neq \varnothing (i \neq j)$, *then* $X_0 \in U/A \wedge X_0 \subseteq A^*(D_i) \wedge X_0 \subseteq A_*(D_j)$ (*suppose that* $D_k = \{x | x \in U, f(x, D) = d_k\} (k = 1, 2, \ldots, h)$). *Here must exist* $x \in X_0$, *then* $f(x, D) = d_i$. *And on the other hand,* $X_0 \subseteq A_*(D_j)$, *then* $f(x, D) = d_j$. *And here* $i \neq j$, *so* $d_i \neq d_j$, *then* $f(x, D) = d_i \neq d_j = f(x, D)$. *At this time the contradiction appears. So above hypothesis is not tenable. And considering the randomicity of i and j, it is concluded that above proposition is tenable.*

Theorem 4 *In Information system, marked with $S = (U, C, D, V, f)$, $\forall A \subseteq C$, here $U/A = \{A_1, A_2, \ldots, A_m\}$ stands for what U is divided by A.*

And here, $U/D = \{D_1, D_2, \ldots, D_h\}$ stands for what U is divided by D. Then here is an equation (shown as following).

$$BND_A(D) = \bigcup_{D_i \in U/D} (A^*(D_i) - A_*(D_i)) = \bigcup_{D_i \in U/D} A^*(D_i) - \bigcup_{D_i \in U/D} A_*(D_i).$$

Refer to Theorem 2 and Theorem 3, above equation can be easily proved.

Theorem 5 *In Information system, marked with $S = (U, C, D, V, f)$, $\forall A \subseteq C$, here $U/A = \{A_1, A_2, \ldots, A_m\}$ stands for what U is divided by A.*

And here, $U/D = \{D_1, D_2, \ldots, D_h\}$ stands for what U is divided by D. Then here is an equation (shown as following).

$$\bigcup_{D_i \in U/D} A^*(D_i) = U$$

The following is probative approach:

It is obvious that the equation, marked with $\bigcup_{D_i \in U/D} A^*(D_i) \subseteq U$, is tenable. Then suppose $\forall x \in U$, $f(x, D) = d_j$, then $[x]_A \subseteq A^*(D_j)$, and $[x]_A \subseteq A^*(D_j)$. Then considering the randomicity of x, the equation, marked with $U/D = \{D_1, D_2, \ldots, D_h\}$, is tenable. And so is the equation, marked with $\bigcup_{D_i \in U/D} A^*(D_i) = U$.

Theorem 6 *In Information system, marked with $S = (U, C, D, V, f)$, $\forall A \subseteq C$, here $U/A = \{A_1, A_2, \ldots, A_m\}$ stands for what U is divided by A.*

And here, $U/D = \{D_1, D_2, \ldots, D_h\}$ stands for what U is divided by D. Then here is an equation (shown as following). $BND_A(D) = U - POS_A(D)$.

Refer to Theorem 4 and Theorem 5, above equation can be easily proved.

Theorem 7 *All the attribute selection standards, include attribute selection standard based on positive region, attribute selection standard based on dependence degree and attribute selection standard based on rough boundary, are equivalence to each other.*

The following is probative approach:

$$|POS_{\{b\}}(D)| = \max_{a \in C}\{|POS_{\{a\}}(D)|\} \overset{\Leftrightarrow \gamma_{\{b\}}(D) = |POS_{\{b\}}(D)|/|U|}{\geq |POS_{\{a\}}(D)|/|U|(\forall a \in C)}$$

$$\Leftrightarrow \gamma_{\{b\}}(D) = \max_{a \in C}\{\gamma_{\{a\}}(D)\} \overset{\Leftrightarrow |BND_{\{b\}}(D)| = |U - POS_{\{b\}}(D)|}{\leq |U - POS_{\{a\}}(D)|(\forall a \in C)}$$

$$\Leftrightarrow |BND_{\{b\}}(D)| = |U - POS_{\{b\}}(D)| = \min_{a \in C}|U - POS_{\{a\}}(D)|$$

Known from above Theorem 7, the three attribute selection standards are equivalence to each other. So it is easy to analyze both advantage and disadvantage of the three attribute selection standard by using attribute selection standard based on positive region.

Known from positive region's definition, it has following disadvantages: (1) When selecting attribute, there are so many attributes who has the same positive region. And it is hard to select the right attribute. If the attribute is selected discretionally, the generated decision tree will be huge. It indicates that this attribute selection standard is quite imprecise. (2) When selecting attribute, it is probably that the selected attribute will generate many leaf nodes. And these nodes include fewer samples. This results in that the generalization ability of these nodes is too small to contain more meanings. Furthermore, it spends much more time in predicting these nodes.

3 Improved Algorithm

Before bringing forward the improved algorithm, a definition should be given.

Definition 5 In Information system, marked with $S = (U, C, D, V, f), \forall a \in C$, here $U/\{a\} = \{A_1, A_2, \ldots, A_m\}$ stands for what U is divided by condition attribute a.

Here, $U/D = \{D_1, D_2, \ldots, D_h\}$ stands for what U is divided by D. Here $POS_A(D) = \underset{D_i \in U/D}{\cup} A_*(D_i)$ stands for positive region that condition attribute set, marked by A, acts on decision attribute set, marked by decision attribute D, in U. Here $|POS_{\{a\}}(D)| + \sum_{i=1}^{h} \frac{|\{a\}_*(D_i)|}{|U|} \frac{|\{a\}_*(D_i)|}{|D_i|}$ is defined as adjoint positive region of condition attribute a. And here it is marked with $BPOS_{\{a\}}(D)$.

Sample 1. $U = \{a1, a2, a3, a4, a5, a6, a7, a8, a9\}$, P, Q and R stands for condition attribute. And Y stands for decision attribute. Here is the following equation:

$$U/Y = \{\{a1, a2, a3, a4, a5, a6\}, \{a7, a8\}, \{a9\}\};$$
$$U/P = \{a1, a2\}, \{a3, a4, a5\}, \{a6, a7, a8, a9\}\};$$
$$U/Q = \{\{a1, a2, a3\}, \{a7, a8\}, \{a4, a5, a6, a9\}\};$$
$$U/R = \{\{a1\}, \{a2\}, \{a3\}, \{a4\}, \{a5, a6, a7, a8\}, \{a9\}\};$$

Known from positive region's definition, the calculation of $BPOS_{\{P\}}(Y)$ is described as following:

$$BPOS_{\{P\}}(Y) = \left|POS_{\{p\}}(Y)\right| + \sum_{i=1}^{h} \frac{|\{P_*\}(Y_i)|}{|U|} \frac{|\{P_*\}(Y_i)|}{|Y_i|} = |\{a1, a2, a3, a4, a5\}|$$

$$+ \frac{|\{a1, a2, a3, a4, a5\}|}{|\{a1, a2, a3, a4, a5, a6, a7, a8, a9\}|} \frac{|\{a1, a2, a3, a4, a5\}|}{|\{a1, a2, a3, a4, a5, a6\}|} = 5 + (5/9)(5/6) = 5.46.$$

With the same calculation approach, the following equations can be easily obtained:

$BPOS_{\{Q\}}(Y) = 5.39; BPOS_{\{R\}}(Y) = 5.41.\ POS_{\{P\}}(Y) = \{a1, a2, a3, a4, a5\};$

$POS_{\{Q\}}(Y) = \{a1, a2, a3, a7, a8\}; POS_{\{R\}}(Y) = \{a1, a2, a3, a4, a9\},$

$|POS_{\{P\}}(Y)| = |POS_{\{Q\}}(Y)| = |POS_{\{R\}}(Y)| = 5; K_P = 5/6; K_Q = 1; K_R = 1.$

If adopting attribute selection standard based on positive region, it is hard to select the three attributes, for their positive region is equal. And if adopting attribute selection standard based on dependence degree, attribute Q or R will be selected, for both of their dependence degree equals 1. But the dependence degree of R can't show its priority well. After all, it is generated by the third category decision value. It is said that R is generated by $Y_3 = \{a9\}$. And with this new kind of attribute selection standard, it sounds reasonable that attribute P will be selected.

Theorem 8 *In Information system, marked with $S = (U, C, D, V, f), \forall a \in C$, here $U/\{a\} = \{A_1, A_2, \ldots, A_m\}$ stands for what U is divided by condition attribute a.*

Here, $U/D = \{D_1, D_2, \ldots, D_h\}$ stands for what U is divided by decision attribute D. And then here is an in equation, marked with $\sum_{i=1}^{h} tfrac|\{a\}_*(D_i)||U| \frac{|\{a\}_*(D_i)|}{|D_i|} \leq 1$. The following is probative approach:

Known from lower approximation's definition, here is following expression:

$\{a\}*(D_i) \subseteq D_i, \overset{h}{\underset{i=1}{\cup}} D_i = U \wedge D_i \cap D_j = \varnothing (i \neq j)$. And then the rational appro-ach, marked with $\sum_{i=1}^{h} \frac{|\{a\}_*(D_i)|}{|U|} \frac{|\{a\}_*(D_i)|}{|D_i|} \leq \sum_{i=1}^{h} \frac{|\{a\}_*(D_i)|}{|U|} \leq \sum_{i=1}^{h} \frac{|D_i|}{|U|} \leq \frac{|U|}{|U|} = 1$ is tenable. Then the probative approach is complete.

Known from Theorem 8, those attributes whose positive region's value is larger than any other are preferentially considered, by using attribute selection standard based on adjoint positive region. If two random attributes' positive region value is not equal to each other, there must be 1 differentiation between them. And if attribute's positive region value is equal to each other, both the number of leaf nodes divided by attribute and the number of sample covered by leaf nodes will be considered posteriorly (here if the leaf nodes divided from selected attribute is same as any other, then these nodes will be merged). This can guarantee that these generated nodes have high generalization ability. In sample 1, although positive region value of the three condition attributes are same to any other, attribute P only generates one leaf node. And this node covers 5 samples. Attribute R generates two leaf nodes. One covers 4 samples, and the other covers 1 sample. Attribute Q generates two leaf nodes. One covers 3 samples, and the other covers 2 sample. According to new attribute selection standard, selecting P is reasonable. Then a new decision tree algorithm, based on adjoint positive region, is proposed in next section. The improved algorithm of generating decision tree, based on adjoint positive region, is marked with **BPOSDT(U,C,D)**.

  BPOSDT(U,C,D)
  Input: training sample set U, condition attribute set C, decision attribute set D
  Output: one decision tree

Algorithm description:

(1) Generate one node marked with N;
(2) If the values that U acts to one attribute are all same, then the node will be marked with this attribute's value. And then return.
(3) If C is null, then the node will be marked with the value of the decision attribute, which is in the majority of training sample set. And then return.
(4) Here R stands for the condition attribute related to $\max_{a \in C} BPOS_{\{a\}}(D)$. And $U/\{R\}$ is equal to $\{U_1, U_1, \ldots, U_k\}$. Here the value of the condition attribute related to $U_i (i = 1, 2, \ldots, k)$ is marked with ui.
(5) Return one decision tree. Its root is marked with R, and its arcs are respectively marked with $u1$, $u2$,…, um. And the sub-trees related to these arcs are respectively marked with NRDT(U_1,C-{R},D, λ), NRDT(U_2,C-{R},D, λ), …, NRDT(U_k,C-{R},D, λ).
(6) In the leaf nodes divided by the same condition attribute, if there exists the nodes whose decision attribute's value is equal, then these nodes should be merged. Here the acts related to this merged nodes are marked with disjunctions of the acts which are not merged.

4 Sample

From Table 1, there are following calculating equations:

$U/\{a1\} = \{\{1, 2, 3\}, \{4, 5, 6, 7\}, \{8, 9\}, \{10, 11, 12, 13, 14, 15\}\};$

$U/\{D\} = \{\{1, 2, 3, 4, 5, 6\}, \{7, 8, 9, 10, 12\}, \{11, 13, 14, 15\}\};$

$$BPOS_{\{a1\}}(D) = |POS_{\{a1\}}(D)| + \sum_{i=1}^{h} \frac{|\{a1\}_*(D_i)|}{|U|} \frac{|\{a1\}_*(D_i)|}{|D_i|}$$

$$= |\{1, 2, 3\} \cup \{8, 9\}| + \frac{|\{1, 2, 3\}|}{|\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15\}|} \frac{|\{1, 2, 3\}|}{|\{1, 2, 3, 4, 5, 6\}|}$$

$$+ \frac{|\{7, 8\}|}{|\{1, 2, \ldots, 15\}|} \frac{|\{7, 8\}|}{|\{7, 8, 9, 10, 12\}|} = 5 + \frac{3}{15} \times \frac{3}{6} + \frac{2}{15} \times \frac{2}{5} = 5 + 0.15 = 5.15$$

$$BPOS_{\{a2\}}(D) = 5.28; BPOS_{\{a3\}}(D) = 5.19; BPOS_{\{a1\}}(D) = 3.15.$$

Since $BPOS_{\{a2\}}(D)$ is largest, a2 should be selected.

$$U/\{a2\} = \{\{1, 2, 3, 4, 5\}, \{6, 7, 8\}, \{9, 10, 11\}, \{12, 13, 14, 15\}\}.$$

And according to above calculating approach, the data set, such as {6,7,8}, {9,10,11} and {12,13,14,15}, should be dealt. Select related condition attribute, and then generate one decision tree shown as following (Fig. 1):

If adopting attribute selection standard, the condition attribute, which is selected firstly, probably is not a2. And this results in that the generated decision

Table 1 Decision table

	a1	a2	a3	a4	D
1	0	1	0	0	0
2	0	1	0	1	0
3	0	1	1	0	0
4	1	1	1	0	0
5	1	1	1	1	0
6	1	2	1	1	0
7	1	2	1	0	1
8	2	2	2	0	1
9	2	3	2	2	1
10	3	3	2	2	1
11	3	3	2	0	2
12	3	4	2	2	1
13	3	4	3	2	2
14	3	4	3	2	2
15	3	4	3	2	2

Fig. 1 Decision tree generated by new algorithm
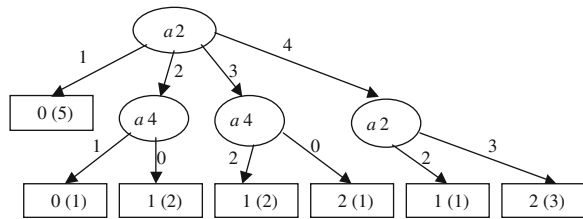
Fig. 2 Decision tree algorithm base on positive region (a1 is selected)
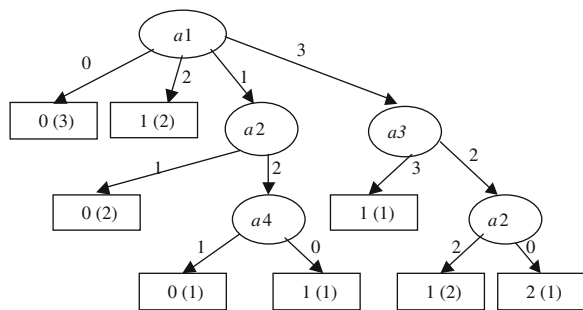
tree not only has much more leaf nodes but also has much deeper leaf nodes. For example, if condition attribute $a1$ or $a3$ is selected at first time, then the generated decision tree is shown as graph 2 and 3 respectively.

Fig. 3 Decision tree algorithm base on positive region (a3 is selected)
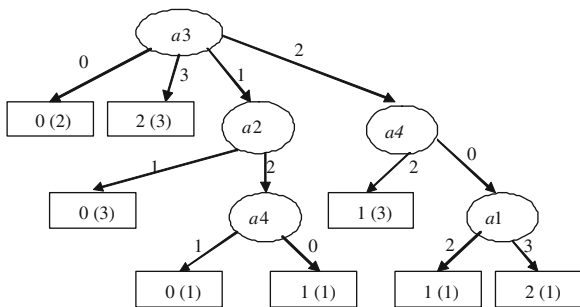
5 Conclusion

In this paper, common attribute selection standards, based on positive region, based on rough bound and based on attribute dependency degree are firstly analyzed in details. At the same time, it is proved that the three kinds of selected attribute standards are equivalent to each other. Then, by taking attribute selection standard of positive region as delegate, the advantages and disadvantages of this algorithm for generating decision tree based on positive region are analyzed. In order to get rid of these disadvantages, a new standard in selecting attributes is proposed in this paper. This attribute selection standard is based on adjoint positive region. And this new standard has following advantages: (1) This standard inherits the attribute selection standard's advantage based on positive region. And the attribute's size of positive region is preferentially considered, which can guarantee that leaf node has less depth in generated decision tree. (2) Only when the attribute's size of positive region is equal to each other, the number of leaf nodes generated from attribute and the number of samples covered by leaf nodes will be taken into account. It is assured that the decision tree has few leaf nodes and its leaf nodes have high generalization ability. The decision tree, generated by using this new standard in selecting attribute, has following advantages: The number of leaf nodes is fewer. And the mean depth of leaf nodes is smaller as well as that the leaf nodes have higher generalization ability. At last, this new attribute selection standard's superiority is described by one sample in details.

References

1. Quinlan JR (1986) Induction of decision trees. Mach Learn 1:81–106
2. Pawlak Z (1982) Rough sets. Int J Comput Inform Sci 11:341–356
3. Wang Guoyin (2001) Rough set theory and knowledge acquisition. Xi An JiaoTong University Press, Xi'an
4. Zhang W (2003) Rough set theory and method. Science press, Beijing
5. Wei J (2003) Rough set based approach to selection of node. Int J Comput Cogn 1(2):25–40
6. Wang M, Wei J (2005) A new decision tree pruning method based on RST. J Northeast Normal Univ 37(3):28–32 (Natural Science Edition)
7. Hong Z (2003) The Research of the simplest decision tree production algorithm based on rough set. J Comp Eng Appl 41(13):129–131
8. Zhao X, Xiang Y, Liu T, Qi Y (2005) An algorithm for decision tree construction based in rough sets. J East China Shipbuild Inst 19(4):73–76 (Natural Science Edition)
9. Guan H, Tian D (2004) Rule abstracting algorithm by decisive tree based on the importance of attribute. Syst Eng Electron 26(3):334–337
10. Chang Z (2005) A method based on rough set to construct decision tree. J Najing Univ Technol 27(4):80–83

A Historical Study About the Developing Process of the Classical Linear Time Series Models

Shu-yuan Nie and Xin-qian Wu

Abstract Through investigating the original literatures of some statisticians who have made key contributions to the development of the ARMA model, the author not only analyzes the evolution process of AR, MA and ARMA model, but also emphasizes the inheritance relation between them. It will lay a foundation and provide a clear clue for the study of the discipline history of time series analysis.

Keywords The history of the ARMA model · AR(2) model · MA model · Correlation · Time domain analysis method

1 Introduction

In present information age, as the commonly used tool for data analysis and processing, time series analysis is drawing scientists' great attention. It has two principle methods, i.e. the frequency domain method and the time domain method. The initial and the most important study objective of the latter was the classical linear ARMA model, which played the role linking the preceding and the following in the history of time series analysis. The ARMA model was connected by the periodogram method, which was the core of the frequency domain analysis method, and the non-stationary ARIMA model, which was also the key part of the time domain analysis method [1]. In spite of its wide application, research literature on its historical development for such a frontier content is still wanting.

S. Nie (✉)
Department of Mathematics, Luoyang Normal University, Luoyang 471022, China
e-mail: nsynsy2008@yahoo.cn

X. Wu
School of Mathematics and Statistics, Henan University of Science and Technology, Luoyang 471023, China
e-mail: wuxinqian1001@163.com

Most of historical introduction in existence are concisely scattered in various monographs and papers [2]. Systematic studies on the historical evolution of the classical linear time series models are therefore necessary and of both theoretical and practical significance.

2 The Background of the Time Domain Analysis Method

It's well-known that the periodogram method was first constructed by Arthur Schuster (1851–1934) and used to solve the periodicity of sun-spot in 1906, which gradually became the main implement of searching for hidden periodicity in all kinds of natural phenomena [3, 4]. But its defects should not be ignored, such as, the periodogram method supposed that time series might contain strict periodicity and the terms of the series were independent. The most terrible thing was that the periodogram method wasn't concerned about the external causal factors [4].

Two main lines were almost synchronized to discuss and modify these problems, both of which originated from George Udny Yule (1871–1951) [5, 6]. One approach was the personal work of Yule who established the autoregression model in 1927 [7]. On the other hand, based on the work of Yule in 1921 and 1926, Evgeny Evgenievich Slutzky (1880–1948) finally established the moving summation model in 1927 [8]. The same features were that the random components were believed to play a positive role and regarded as an essential part of time series in the above models. It is well known that Yule and Slutzky commonly carried out the revolution of statistics by seeing the error term as the accidental external random disturbance. Therefore, when the periodogram method was met with a severe criticism in many fields, these new ideas began to be widely used, which gradually induced time series analysis to the era of the time domain analysis method [9].

3 The Establishing Process and the Inheritance Relationship of the ARMA Model

The ARMA model can be particularly divided into three models: autoregression model (AR model), moving average model (MA model) and autoregression moving average model (ARMA model). Two statisticians who have made important contributions to the AR model were Yule and Gilbert Thomas Walker (1868–1958).

3.1 The AR(2) and AR(4) Models First Established by Yule

The study area of Yule was quite wide [10]. All of his three key papers about time series were published in 1920s. It was quite interesting that Yule began to explore the autoregression problem because of his puzzle why statisticians often got the nonsense correlations from time series data.

In 1921, Yule first generalized and reviewed some famous studies, including the study on correlations made by John Henry Poynting (1852–1914), Reginald Hawthorn Hooker (1867–1944), Cave-Browne-cave, F. E. (1865–1943) and Monsieur March (1859–1933) etc. He especially emphasized the high-difference method advocated by William Sealy Gosset (1876–1936) and Oskar Anderson (1887–1960), who even assumed variables were the sum of the random residuals and a polynomial function of the time. Yule brought up his suspicion to the rationality of separating the random residuals and doing the nth difference [5]. By analyzing the variate-difference method with the harmonic function, Yule had drawn some conclusions. But he couldn't accept the assumption that variables were functions of time and was confused with the time-correlation problem. So combining with his physics background, Yule further explored the question of the nonsense correlation. In his second paper, Yule clearly explained the implication of the nonsense correlation with a practical instance and classified time series into three types by the nature of their serial correlation [6]:

(a) Random series, a, b, c, d...
(b) Conjunct series the differences of which are random, a, a + b, a + b + c...
(c) Conjunct series the differences of which are themselves conjunct series, a, 2a + b, 3a + 2b + c...

It was in the discussion of classification that Yule found the distribution of correlations especially depended on the correlation of the first difference, not the correlation of the series. It was the third type series that tended to yield high correlation, whether the true value of the correlation existed or not. That was to say, the third type series was the dangerous series which would mislead statisticians to get the wrong conclusions.

After he could clearly identify and describe the nonsense correlation series from the angle of mathematics, Yule confirmed that the series had their own life clock and at least were partly self-determined. So Yule tried to explore the high correlation would be generated among which intervals, stations and previous values, which was his instructive thought to be used for constructing autoregression analysis.

When he analyzed the pendulum and Wolfer's sunspot numbers with harmonic functions in 1927, Yule first concluded if a disturbance ε occurred, the general term u_x and its two lag terms u_{x-1}, u_{x-2} had such a relation [7]:

$$u_x = (2 - \mu)u_{x-1} - u_{x-2} + \varepsilon$$

The formulation was famous AR(2) model. It was the first example in which the term of time series was determined by its lag terms and term ε was not regarded as error of observation but accidental disturbance. This model was deemed the foundation and origin of modern time series analysis [11].

When Yule assumed the harmonic function had two periods, he attained AR(4) model [7]:

$$u_x = k_1(u_{x-1} + u_{x-3}) - k_2 u_{x-2} - u_{x-4} + \varepsilon$$

More lag terms (from u_{x-1} to u_{x-4}) had been applied to estimate u_x in this formulation, which should be more advanced than AR(2) model. But the result was worse and not better. The AR(4) model wasn't appreciated by the present statisticians because it had confined a particular form of the regression equation. For example, the coefficients of u_{x-1} and u_{x-3} were identical, and the coefficient of u_{x-4} was unit.

Yule also analyzed the more general method of regression equation $u_x = b_1 u_{x-1} - b_2 u_{x-2}$ and proved with two experiments that the periodogram method should be replaced by the regression analysis method. At last, Yule put forward his innovation thoughts that not all the time series are regarded as the functions of time, and there exist such time series related to its lag variables [7]. Therefore, the AR models first established by Yule objectively reflected the relation of the variable in time series. And it was also an amendment and improvement to commonly regard time series as the functions of time. Yule was fully deserved the creator of modern time series analysis.

3.2 The AR(s) Model Expanded by Walker

In 1931, another English statistician Walker expanded Yule's conclusion and gave the more general AR(s) model [12]:

$$u_x = g_1 u_{x-1} + g_2 u_{x-2} + \cdots + g_s u_{x-s} + v_x$$

where v_x represent disturbance. For Walker, the thought and the inferring process of establishing this model was rather simple. His focus achievement was he successfully applied this model to solve the practical problem. For example, Walker concluded if series u had a relation

$$u_x = g_1 u_{x-1} + g_2 u_{x-2} + \cdots + g_s u_{x-s}$$

When the number of observations was great, autocorrelation coefficient series r met with a similar relation

$$r_y = g_1 r_{y-1} + g_2 r_{y-2} + \cdots + g_s r_{y-s}$$

Thus two series must have the same periods. But series r was based on the whole series, it was much less influenced by accidental effects and its figure was

more smooth than the figure of series u. Therefore, it was more precise to identify the characteristic of series u using the value of r than using that of u. So, Walker took autocorrelation coefficient series r and correlation-periodogram as his essential tool in the whole study. Furthermore, he discussed the relationship between series r and the Fourier terms of series u and drew some conclusions. For example, Walker proved that the relation of the amplitude ratio f_q and $f_q^{/}$ of two series u and v was

$$\frac{f_q^2}{f_q^{/2}} = \frac{1 - r_1^2}{1 - 2r_1 \cos q\alpha + r_1^2}$$

where r_1 was the correlation coefficient between consecutive terms of series u, q denoted interval, $\alpha = \frac{2\pi}{n}$. Walker also concluded that the correlation coefficient r_s between terms of series u separated by s interval met with

$$r_s = f_1^2 \cos s\alpha + f_2^2 \cos 2s\alpha + \cdots + f_m^2 \cos ms\alpha.$$

Walker used these formula to explain the pressure problem at Port Darwin in two ways. At last, he calculated some tables of correlation coefficients and correlation-periodogram and used his method to determine whether the oscillations were due to the fact that the external disturbances were cyclical or the randomly disturbed system itself had an inherent damped nature [11, 12]. The study about the damp oscillation and the thought of identifying the stationary random process eventually lead Walker and Yule to establish the auto regression model, which became an ideal tool little by little and widely used in ARMA and ARIMA and VAR models. Walker's contribution should also be paid great attention.

3.3 The MA(n) Model Constructed by Slutzky

Being different from other statisticians, Slutzky was more interested in the random disturbances. He definitely pointed out that the terms of an empirical series were not independent but sometimes correlated closely. It was obviously a direct negation to the idea of the periodogram method which regarded the terms of series as independent items. With a brief English summary, his paper was first published in Russian in 1927 and the whole article was translated into English and published again 10 years later [8].

The work of Slutzky was based on the study of Yule, who concluded in 1921 and 1926 that after a fixed order differences to a purely random series, the series would show a tendency to regular fluctuations [5, 6]. Slutzky tried to explore "is it possible that a definite structure of a connection between random fluctuations could form them into a system of more or less regular waves [8]?" To solve the problem, Slutzky supposed that the consequences y were determined by the different causes x and gave different weights A according to the influence of the causes. He gave many specific models, such as,

$$y_i = A_0 x_i + A_1 x_{i-1} + \cdots + A_{n-1} x_{i-(n-1)}$$

Just through discussing and modifying the above models, Slutzky finally established the model of the cumulative effects of random disturbances [8]:

$$\xi(t) = b_0 \eta(t) + b_1 \eta(t-1) + \cdots + b_n \eta(t-n)$$

where $\eta(t)$ was the random disturbances, $\eta(t-1), \ldots, \eta(t-n)$ denoted the lag terms of $\eta(t)$. Slutzky perhaps more emphasized "summation". He called it the moving summation model, which was named "the moving average model" by Wold in 1938 [13]. Slutzky not only compared his model with Dorothy Swaine Thomas's index of English business cycles [8], but also established more sophisticated models and discussed their correlational functions. At last, he proved the law of the sinusoidal limit—Repeating indefinitely moving summations followed by repeated differencing of a random series produced a sinusoidal process [8, 11].

It was the theorem that made Slutzky to be closely contacted with Yule. Nowadays the Yule-Slutzky effect was well-known. Certainly, Slutzky's MA model had been used to explain many phenomena for the use of moving averages of random causes.

3.4 The ARMA(s, n) Model Established by Wold

The last core character was Herman Wold (1908–1992), who was the famous econometrician and statistician in Sweden and had done lots of fundamental study on diverse disciplines. In his writing in 1938, Wold clearly indicated that his study objective was the discrete stationary time series [13]. He also proposed the following ideas:

The foundation of the periodogram method was rather narrow, but the hypothesis of hidden periodicity needn't to be separately treated. The AR and MA model, including the periodogram method could be merged into the process of linear regression. But from the viewpoint of the theory of probability, the process of linear regression was a special case of the stationary random process defined and studied by Aleksandr Yakovlevich Khinchin (1894–1959). Therefore, Wold first integrated the empirical investigations with the probability theory of the stationary random process. The main characters of the former were Schuster, Yule, Slutzky, Walker and so on, while the latter was Khinchin.

Wold analyzed some types of cases of the discrete stationary process in detail, such as,

(a) the purely random process

$$F(t_1, t_2, \ldots, t_n, u_1, u_2, \ldots, u_n) = F(u_1)F(u_2)\cdots F(u_n)$$

(b) the process of moving averages

$$\xi(t) = b_0\eta(t) + b_1\eta(t-1) + \cdots + b_h\eta(t-h)$$

(c) the general process of linear regression

$$\{\xi(t)\} = b_0\{\eta(t)\} + b_1\{\eta(t-1)\} + b_2\{\eta(t-2)\} + \cdots$$

(d) the process of linear autoregression

$$\xi(t) = \eta(t) - a_1\xi(t-1) - a_2\xi(t-2) - \cdots - a_h\xi(t-h)$$

(e) the periodic processes

$$\xi(t) - \xi(t-h) = 0$$

(f) the process of hidden periodicities

$$\{\xi(t)\} = \left\{\xi^{(1)}(t)\right\} + \left\{\xi^{(2)}(t)\right\} + \cdots \left\{\xi^{(k)}(t)\right\}$$

Wold merged these models into the general stationary process and analyzed its structural properties from an unexplored angle. He proved the famous decomposition theorem, which was also called the orthogonal decomposition theorem [13]. The decomposition theorem was the soul of the modern time series analysis, which revealed the conversion relationship among the different models, such as, the AR model, MA model and ARMA model. It was based on this theorem that Wold finally established the ARMA(s, n) model:

$$\xi(t) = a_1\xi(t-1) + a_2\xi(t-2) + \cdots + a_s\xi(t-s) + \eta(t) + b_1\eta(t-1) + b_2\eta$$
$$(t-2) + \cdots + b_n\eta(t-n)$$

where $\eta(t)$ was also the random disturbances, $\xi(t-1), \ldots, \xi(t-s)$ and $\eta(t-1)$, $\ldots, \eta(t-n)$ denoted the lag terms of $\xi(t)$ and $\eta(t)$. It was obviously that Wold's comprehensive study was of great significance [14]. He not only strictly demonstrated that the models of hidden periodicities, linear autoregression and moving averages were all special cases of the discrete stationary process, but also further compared the correlograms of these important models and used them to analyze some specific series. Furthermore, Wold's comprehensive study can be generalized in many directions. For example, the Wold decomposition theorem was generalized to the non-stationary case by his teacher Harald Cramér (1893–1985). Wold's simple regression analysis was popularized to the multi-dimensional case by his pupil Peter Whittle (1927–). Last, ARMA model laid the groundwork for the ARIMA model established by George Edward Pelham Box (1919–) and Gwilym Meirion Jenkins (1933–1982) in 1970s [15]. Wold was a key figure in the developing process of the stationary time series.
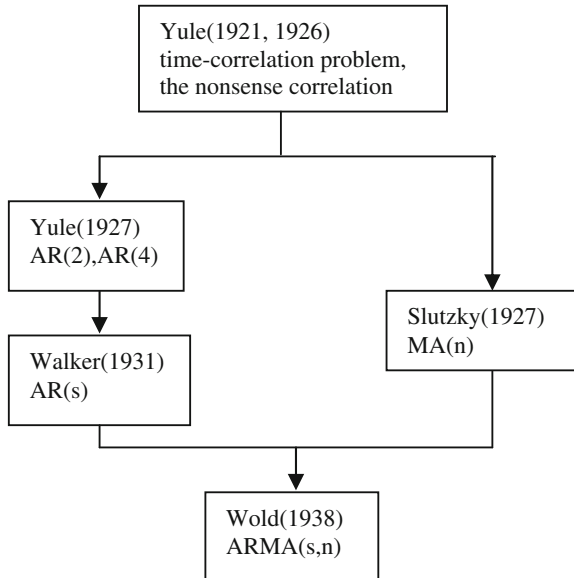
4 Conclusion

From the above study, it is easy to draw the conclusion. Based on the study of Yule in 1921 and 1926, Yule and Slutzky respectively established the AR and MA model in 1927. Walker gave the AR(s) model in 1931. It was Wold who finally summarized all the thoughts and established the ARMA model in 1938. The historical evolution process and their inheritance relationship of three stationary models is as follows:

The paper has laid a foundation and provided a clear clue for the discipline history study of time series analysis.

References

1. Brockwell PJ, Davis RA (2002) Introduction to time series and forecasting, 2nd edn. Springer, New York
2. Chen Y-H (2010) Semiparametric marginal regression analysis for dependent competing risks under an assumed copula. J Roy Stat Soc B 72B(2):235–257
3. Schuster A (1906) On sun-spot periodicities. Preliminary notice. Proc Roy Soc Lond. Ser A 77(515):141–145

4. Schuster A (1906) On the periodicities of sunspots. Philos Trans Roy Soc Lond Ser A 206:69–100
5. Yule GU (1921) On the time-correlation problem, with especial reference to the variate-difference correlation method. J Roy Stat 84(4):497–537(1921)
6. Yule GU (1926) Why do we sometimes get nonsense-correlations between time-series?—a study in sampling and the nature of time-series. J Roy Stat 89(1):1–63
7. Yule GU (1927) On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. Philos Trans Royal Soc Lond Ser A 226:267–298
8. Slutzky EE (1937) The summation of random causes as the source of cyclic processes. Econometrica 5(2):105–146
9. Kirchgässner G, Wolters J (2007) Introduction to modern time series analysis. Springer, New York
10. Kendall MG (1952) Review of the international statistical institute: George Udny Yule, 1871–1951. 20(1):92–93
11. Klein JL (1997) Statistical visions in time: a history of time series analysis, 1662–1938. Cambridge University Press, Cambridge
12. Walker G (1931) On periodicity in series of related terms. Proc Roy Soc Ser A 131(818):518–532
13. Wold H (1954) A study in the analysis of stationary time series. Almqvist and Wiksell, Stockholm
14. Whittle P (1992) Obituary: professor Herman wold. J Roy Stat Soc Ser A 155(3):466–469
15. Box GEP, Jenkins GM (1970) Time-series analysis: forecasting and control. Holden Day, San Francisco

A Novel Attributes Partition Method for Decision Tree

Zhen Li, Aili Han and Feilin Han

Abstract In the decision tree's making phase, it is frequent to find the optimal partition of elements with different values of a category attribute at a node. This needs to search over all the partitions for the one with the minimal impurity, which is exponential in n. We present a new heuristic search algorithm, SORT_DP, to find an effective partition, which is polynomial in n. The method uses the mapping from the class probability space to the sub-spaces and the technique of dynamic programming. By comparing the performance against other methods through experiments, we demonstrated the effectiveness of the new method.

Keywords Intelligent computing · Decision tree · Dynamic programming · Optimal partition

1 Introduction

Decision trees are widely used in data mining applications, and are usually fast compared with other classification methods. They have the advantages of easily converting to rules and SQL queries. Especially, a binary decision tree is easy to implement, so it is frequently used in many methods, such as SLIQ. When generating a decision tree, the main step is to determine which attribute is used to

Z. Li · A. Han (✉)
Department of Computer Science, Shandong University, Weihai 264209, China
e-mail: hanal@sdu.edu.cn

Z. Li
e-mail: lizhensd1988@gmail.com

F. Han
College of Software, Shandong University, Jinan 250101, China
e-mail: hanfl9253@163.com

partition at each node. The aim is to find the optimal partition that minimizes the impurity of each node or maximizes the impurity improvement.

There are many kinds of impurity measures, two of them are the GINI index and entropy. Let $\boldsymbol{n} = (n_1,\ldots,n_k)$ be a vector of nonnegative integers representing the number of elements in each class, and let $N = \sum_{i=1}^{k} n_i$ be the total number of elements at each node. The GINI index $g(n) = 1 - \sum_{i=1}^{k} n_i^2/N^2$, and the weighted GINI index $G(n) = Ng(n) = \sum_{i \neq j} n_i n_j/N$. The entropy $h(n) = -\sum_{i=1}^{k} n_i/\log n_i/N$, and the weighted entropy $H(n) = Nh(n) = -\sum_{i=1}^{k} n_i \log n_i$.

We use $I(n)$ to denote the general impurity measure instead of $G(n)$ or $H(n)$. As we all know, the impurity of a node is never less than the sum of the impurities of its children nodes. Hence, the impurity measures are concave functions. Let $\boldsymbol{A}$ be an index set representing the values of the given category attribute. For each attribute value $\alpha \in \boldsymbol{A}$, the corresponding vector is $\boldsymbol{n^\alpha} = (n_1^\alpha,\ldots,n_k^\alpha)$, where n_i^α is the number of elements in class i in which the given attribute takes the value.

Let $\pi = \{B_1, B_2\}$ be a 2-way partition of $\boldsymbol{A}$. It partitions the N elements into two parts: the number of elements in part 1 is $\boldsymbol{N}^1 = \sum_{\alpha \in B_1} \boldsymbol{n^\alpha} = (N_1^1, \ldots, N_k^1)$ the number of elements in part 2 is $\boldsymbol{N}^2 = \sum_{\alpha \in B2} \boldsymbol{n^\alpha} = (N_1^2, \ldots, N_k^2)$. The optimal partition minimizes the sum of the impurities $\sum_{l=1}^{2} I(\boldsymbol{N}^l)$. For each attribute value $\alpha \in \boldsymbol{A}$, the vector of class probabilities is $\boldsymbol{p^\alpha} = (p_1^\alpha, \ldots, p_k^\alpha)$, where $p_i^\alpha = n_i^\alpha / \sum_{j=1}^{k} n_j^\alpha$. The $\boldsymbol{p^\alpha}$ lies in a $(k-1)$-dimension probability space $\omega_{k-1} = \{(x_1,\ldots,x_k)|x_i \geq 0, \sum x_i = 1\}$. Any partition of ω_{k-1} into S_1 and S_2 results in a partition of $\boldsymbol{A}$ into B_1 and B_2, where $B_1 = \{\alpha|\boldsymbol{p}^a \in S_1\}$ and $B_2 = \{\alpha|\boldsymbol{p}^a \in S_2\}$.

In this paper, we discuss the problem of finding the optimal partition of elements for a category attribute with n values in a k-class binary decision tree. The problem can also be described as that of finding optimal partition of putting n elements into 2 sets. A partition of n values for an attribute results in a partition of the corresponding elements. Based on the value of attribute, each element is put into a set. The number of partitions of n values is exponential in n, e.g., there are $2^{n-1}-1$ partitions for a binary tree. We present a new heuristic search algorithm, SORT_DP, to find an effective partition, which is polynomial in n. By comparing the performance against other methods, we demonstrated the effectiveness of the new method.

2 Related Researches

There are previous theories and methods on the optimal partition problem, some of them have good results. For a two-class binary tree, $k = 2$, Breiman et al. [1] proved that the optimal split, under an arbitrary concave impurity measure, satisfies the condition that there exists a threshold p' such that all the attribute values

α of satisfying $p_1\alpha \leq p'$ belong to one class and other values belong to the other class. It means that searching for the optimal partition is linear in n. Chou [2] proved that, for a general impurity measure, the necessary condition for the optimal partition is follows: each set must satisfy the principle of being nearest from the centroids. For a binary tree, this condition can be reduced to the existence of a hyper-plane in the class probability space. Chou gave a heuristic method based on a K-means clustering algorithm, but did not give the conclusion on the performance. Burshtein et al. [3] presented a different condition satisfied by the optimal partition for a general concave impurity measures. Especially, for a binary tree, it can be reduced to the partition by a single hyper-plane in the class probability space. Based on this, they designed an algorithm of enumerating all the linear partitions.

These theories do not offer a practical algorithm, but some simple heuristic algorithms have better performance. The Flip-Flop method by Nadas et al. [4] uses the idea from the special case of two-class binary trees to restrict the set of partition. The method starts with a randomly binary partition of the attribute values, and then it reverses the role of attribute and class, obtaining a two-"class" problem in which the attribute values have been partitioned. In the third step, it reverses the role of attribute and class again. Repeat steps 2 and 3 until the resulting partition does not change. Another heuristic search algorithm is the SLIQ [5] method, in which Metha et al. cited IND [6] as the initial source where the method was proposed. For $n \leq 10$, the SLIQ method searches over all $2^{n-1}-1$ 2-way partitions; for larger n, it uses a greedy algorithm that starts with an empty bin and a full bin. It adds one element to the first bin which gives the best split. Repeat the process until the second bin is empty. The computational complexity of the SLIQ algorithm is $O(n^2)$. Coppersmith et al. [7] summarized the ideas mentioned above and proposed a new heuristic algorithm, PC. The main idea of PC is to select a particular direction in the class probability space and then consider the hyper-plane that are perpendicular to this direction. They assume that the first principal component of $\boldsymbol{p}^\alpha$, $\alpha \in \boldsymbol{A}$, is a good choice. In addition, there exists other approaches, such as QUEST [8]. The QUEST method uses the discriminatory information to define the space and order the transformed values.

3 Algorithm for Attributes Partition

By Chou's inspiration, the problem of finding an optimal partition can be regarded as the clustering problem in a class probability space. Since the class probability space is $(k-1)$-dimension, this problem turns into a high-dimension clustering problem when k is large. Subspace clustering [9] is an effective method for a high-dimension clustering problem, which is an extension of feature selection that attempts to find clusters in different subspaces of the same dataset.

Definition

1. The relative weight of class C_i in the vector $\boldsymbol{p}^\alpha$, $\alpha \in A$, is defined as follows:

$$R(C_i, \alpha) = n_i^\alpha / \sum_{j=1, j\neq i}^{k} n_j^\alpha \tag{1}$$

Consider the special case, a two-class binary tree. We sort the attribute values $\alpha \in \mathbf{A}$ by $R(C_i, \alpha)$ to generate a sequence $A' = \langle \alpha_1, \ldots, \alpha_n \rangle$, which satisfies the condition that the optimal partition must appear between α_i and α_{i+1}, $1 \leq i < n$. Since it requires only $n-1$ computations of the impurity, the search for the optimal partition is linear in n. For a two-class binary tree, we map the class probability space to a 1-dimension subspace by means of the relative weight $R(C_i, \alpha)$, and then search the optimal partition in the subspace. The method for a two-class binary tree can be generalized for a k-class binary tree. By enumerating class $C_i, i \in [1, k]$, we map the class probability space to the $(k-1)$-dimension subspace determined by class C_i by means of the relative weight $R(C_i, \alpha)$, and then search the optimal partition in the subspace. In practice, searching in the $(k-1)$-dimension subspace determined by class C_i means that C_i is the main factor of affecting the partition, so the optimal partition is decided by $\boldsymbol{n}_i$. In other words, the optimal partition can be found in polynomial time of n in the subspace determined by class $C_i, 1 \leq i < n$.

3.1 Optimal Substructure

We assume that class $C_i, i \in [1, k]$ is the main factor of affecting the partition. Sort the attribute values $\alpha \in \boldsymbol{A}$ by the relative weight $R(C_i, \alpha)$ to generate a sequence $A' = \langle \alpha_1, \ldots, \alpha_n \rangle$, where $R(C_i, \alpha_1) \geq R(C_i, \alpha_2) \geq \cdots \geq R(C_i, \alpha_n)$. For the sake of convenience, let B_1 be the set of attribute values generated by the optimal partition, that is, B_1 is generated by the optimal partition of considering the whole sequence $A' = \langle \alpha_1, \ldots, \alpha_n \rangle$. We define B_1^i be generated by the optimal partition of considering the partial sequence $\{\alpha_1, \ldots, \alpha_i\}$, that is, $B_1^i \subseteq \{\alpha_1, \ldots, \alpha_i\}$. According to the assumption, class C_1 is the main factor of affecting the partition, and the sequence A' is sorted by the relative weight of class C_i in decreasing order. This means that B_1^i is included in B_1, that is, the optimal partition of the sub-problem is included in the optimal partition of the problem. Therefore, the attribute partition problem has the property of optimal substructure.

3.2 Recursive Formula

Based on the property of optimal substructure, we use the technique of dynamic programming to search for the optimal partition in the subspace determined by class C_i.

Let $F(i)$, $i \in [1, n]$, be the current optimal partition of satisfying $B_1^i \subseteq \{\alpha_1, \ldots, \alpha_i\}$. $F(i)$ is initialized to be $\{B_1^i, B_2^i\}$, where $B_1^i = \{\alpha_1, \ldots, \alpha_i\}$ and $B_2^i = \{\alpha_{i+1}, \ldots, \alpha_n\}$. Let $B_1^1, B_1^2, \ldots, B_1^i$ be the sequence of the optimal partition of the sub-problems. In order to calculate B_1^{i+1}, we only need to add α_{i+1} to $B_1^1, B_1^2, \ldots, B_1^i$, and then calculate the impurity of the new partitions. The recursive formula is as follows:

$$F(i) = \begin{cases} \{B_1^1, B_2^1\} & \text{if } i = 1 \\ \{B_1^i, B_2^i\} & \text{if } (i > 1) \,\&\, I(F(i)) \leq \min_{1 < i} I(F(j)) \\ \{B_1^j \cup \{\alpha_i\}, B_2^j - \{\alpha_i\}\}, & \text{where} I(\{B_1^j \cup \{\alpha_i\}, B_2^j - \{\alpha_i\}\}) = \min_{1 \leq j < i} I(F(j)) \\ & \text{if } (i > 1) \,\&\, I(F(i)) > \min_{1 \leq j < i} I(F(j)) \end{cases} \quad (2)$$

The optimal partition in the subspace determined by C_i is stored in $F(i)$, $i \in [1, n]$. Enumerate k classes and calculate $F(i)$, $i \in [1, n]$, we can find the partition with the minimal impurity, i.e. the optimal partition.

3.3 Algorithm Description

Algorithm SORT_DP
For $m \leftarrow 1$ to k //k: the number of class values
$A' = \langle \alpha_1, \ldots, \alpha_n \rangle$, from sorting A by $R(C_m, \alpha)$;
For $i \leftarrow 1$ to n //n: the number of attribute values

$\quad B_1^i = \{\alpha_1, \ldots, \alpha_i\}; \quad B_2^i = \{\alpha_{i+1}, \ldots, \alpha_n\}; \quad F(i) \leftarrow \{B_1^i, B_2^i\};$

End for
For $i \leftarrow 2$ to n

$\quad$ For $j \leftarrow 1$ to i-1

$G(j) \leftarrow \{B_1^j \cup \{\alpha_i\}, B_2^j - \{\alpha_i\}\};$
If $I(F(i)) > I(G(j))$ then $F(i) \leftarrow G(j);$

$\quad$ End for

End for
Record $\leftarrow \min_{1 \leq i \leq n} I(F(i));$
End for
Return Record;

The time complexity of the SORT_DP algorithm is analyzed as follows. The inner loop includes two parts: sorting and dynamic programming. The time complexity of sorting is $O(n\log n)$, and that of dynamic programming is $O(n^2)$. Thus, the time complexity of the inner loop is $O(n^2)$. For k classes, the inner loop runs k times, so the time complexity of the SORT_DP algorithm is $O(kn^2)$.

4 Experiments and Discussions

A series of Monte Carlo simulation experiments were run to compare the performance of the methods for the attribute partition problem. We only compare our algorithm with the algorithms in polynomial time of n, such as SLIQ and PC. Each simulation experiment considers a category attribute taking n distinct values and a class attribute taking k distinct values. For each category attribute value α, a vector of the class frequencies $\boldsymbol{n}^{\alpha}$ is generated by selecting each element independently from a set $\{0,\ldots,n-1\}$ with uniform distribution. This makes the vector uniformly distribute over the class probability space.

Each group of simulation experiments was repeated M times, $M = 50{,}000$. Obviously, a good algorithm should satisfy the following conditions: the attribute with the biggest probability is used to produce the optimal partition; if it cannot find the optimal partition, the error should be as small as possible. We designed several statistic methods to measure the performance, including *% Fail*, *Ave*, and *Max*.

Let *Fail* be the number of failing to find the optimal partition, and *%Fail* the ratio of *Fail* to the total number of experiments. The smaller the rate *%Fail*, the better the algorithm. Let *diff* be the difference between the impurities of the sub-optimal partition and the optimal partition, and *error* the quotient of *diff* to the impurity of the optimal partition. The average error *Ave* is the quotient of the sum of *error*s to the number of experiments. The maximal error *Max* is the maximum of *error*s. We use *Ave* and *Max* to analyze the performance when the algorithm can only get the sub-optimal partition.

4.1 Experimental Results and Analysis

The first group of experiments compared the methods SLIQ, PC, and SORT_DP in terms of the ability of finding the optimal partition. The experiments cover a range of n and k, and the results are listed in Table 1 with GINI index or entropy. The large values of n were not taken into account because of the computational complexity. From Table 1, we know that the SORT_DP algorithm performs better than SLIQ and PC, and that SLIQ and PC have better performance on GINI index than on entropy.

For the first group of experiments, we gave the plots of *% Fail* versus *n* from 2 to 30 for the algorithms SLIQ, PC and SORT_DP with GINI, $k = 10$ or 20, $M = 50,000$, which are shown in Fig. 1. From Fig. 1, we know that, as *n* and *k* increase, all the methods become less success at finding the optimal partition, but the deterioration of the SORT_DP algorithm is lower than that of SLIQ and PC.

In the second group of experiments, we introduced the two measures: the average error *Ave* and the maximal error *Max*. The experiments covered a range of *n* and *k* values that are used in the first group. The experimental results are listed in Table 2. From Table 2, we know that the errors of the three methods are all small, and the SORT_DP algorithm has better performance than SLIQ and PC.

For the second group of experiments, we gave the plots of *Ave* versus *n* from 2 to 30 for the algorithms SLIQ, PC and SORT_DP with GINI, k = 10, M = 50,000 and the plots of *Max* Versus *n* from 2 to 30 for the three algorithms with GINI, k = 10, M = 50,000, which are shown in Fig. 2. From Fig. 2, we know that the errors of the three methods are all very small and the SORT_DP method has better performance than the other two methods when using the measure of *Ave* or *Max*.

In the third group of experiments, we select $n = 50$. In this group, we use the best partition from the three methods SLIQ, PC and SORT_DP instead of the optimal partition since we can not search over all the partitions. The results are listed in Table 3, which indicates that the SORT_DP method still has better performance than SLIQ and PC, but the difference is not evident.

Table 1 *% Fail* for the algorithms SLIQ, PC, SORT_DP

n	Algorithm	With Gini index				With entropy			
		$k = 5$	$k = 10$	$k = 15$	$k = 20$	$k = 5$	$k = 10$	$k = 15$	$k = 20$
8	SLIQ	7.158	10.240	11.714	12.302	7.766	10.992	13.506	15.254
	PC	5.434	3.936	3.822	3.924	5.488	7.640	9.224	10.004
	SORT_DP	0.000	0.000	0.000	0.000	0.004	0.002	0.001	0.000
10	SLIQ	11.930	17.012	18.692	19.730	13.528	18.042	22.406	25.162
	PC	10.524	8.942	8.660	8.660	9.088	13.214	15.956	17.534
	SORT_DP	0.018	0.004	0.000	0.000	0.052	0.014	0.002	0.000
12	SLIQ	15.824	22.120	25.082	26.592	16.152	19.158	23.348	26.108
	PC	16.178	13.848	13.872	14.228	12.406	18.668	21.986	23.748
	SORT_DP	0.140	0.040	0.010	0.000	0.160	0.100	0.020	0.004
14	SLIQ	19.728	27.228	30.346	32.172	20.854	28.462	30.338	32.088
	PC	21.634	19.554	19.980	20.288	15.516	20.587	26.948	29.110
	SORT_DP	0.586	0.380	0.144	0.056	0.498	0.246	0.086	0.016
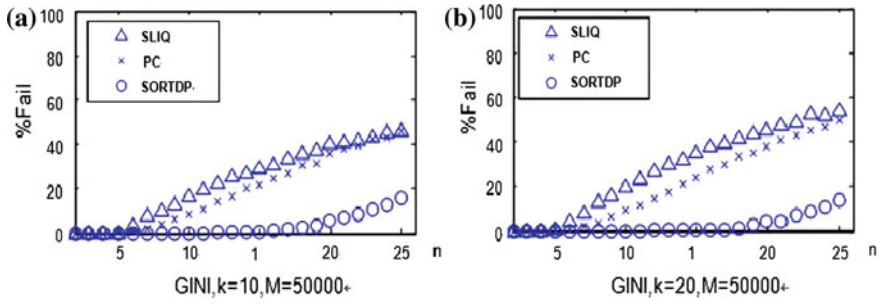
Fig. 1 **a** Plots of % *fail* versus *n* from 2 to 30 for the algorithms SLIQ, PC and SORT_DP with GINI, $k = 10$ and $M = 50{,}000$; **b** Plots of % *fail* versus *n* from 2 to 30 for the algorithms SLIQ, PC and SORT_DP with GINI, $k = 20$ and $M = 50{,}000$

Table 2 % *Fail*, *ave* and *max* for the algorithms SLIQ, PC, SORT_DP

n	*k*	Algorithm	With Gini index			With entropy		
			% Fail	*Ave*	*Max*	*% Fail*	*Ave*	*Max*
8	5	SLIQ	7.158	0.634	39.650	7.766	0.214	27.819
		PC	5.434	0.392	42.864	5.488	0.578	54.525
		SORT_DP	0.000	0.000	0.000	0.004	0.0001	33.880
10	10	SLIQ	17.012	1.221	32.005	18.042	1.051	23.485
		PC	8.942	0.409	24.866	13.214	1.346	51.169
		SORT_DP	0.004	0.0001	2.241	0.014	0.0003	4.793
12	15	SLIQ	25.082	1.611	38.697	23.348	2.013	32.811
		PC	13.872	0.546	26.076	21.986	2.212	48.220
		SORT_DP	0.010	0.0002	5.810	0.020	0.0009	10.207
14	20	SLIQ	32.172	1.8754	31.907	32.088	3.002	46.765
		PC	20.288	0.698	21.676	29.110	2.840	42.371
		SORT_DP	0.056	0.0006	3.116	0.016	0.0005	8.330
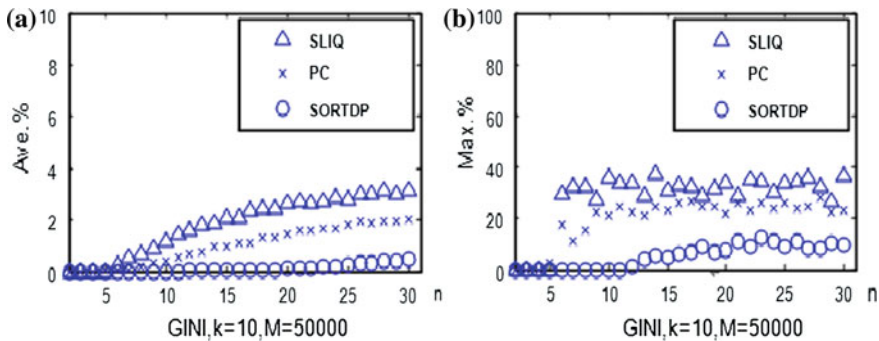
Fig. 2 **a** Plots of *ave* versus *n* from 2 to 30 for the algorithms SLIQ, PC and SORT_DP with GINI, $k = 10$, $M = 50{,}000$; **b** Plots of *max* versus *n* from 2 to 30 for the three algorithms with GINI, $k = 10$, $M = 50{,}000$

Table 3 *% Fail* for the algorithms SLIQ, PC, SORT_DP with $n = 50$

n	k	Algorithm	With Gini index	With entropy
			% Fail	*% Fail*
50	5	SLIQ	49.932	54.144
		PC	71.456	79.600
		SORT_DP	36.964	25.920
50	10	SLIQ	58.870	63.441
		PC	67.088	76.401
		SORT_DP	54.740	42.960
50	15	SLIQ	61.496	64.704
		PC	64.914	72.420
		SORT_DP	60.354	51.220
50	20	SLIQ	62.554	64.560
		PC	64.560	71.000
		SORT_DP	62.538	55.228

5 Conclusions

This paper presents a new heuristic method for partitioning the elements at each node of a binary decision tree when the partition is based on a category attribute. Through comparing with the SLIQ and PC methods, we demonstrate the effectiveness of the new method. Searching for the optimal partition needs to traverse all the $2^{n-1}-1$ partitions, but the time complexity of the SORT_DP method is only $O(kn^2)$, and in the case of getting the sub-optimal partition, the error between the optimal and sub-optimal partitions is very small. This makes the SORT_DP method be practical in real data mining applications in which a binary decision tree is used as the classification model.

References

1. Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and regression trees. Wadsworth International Group, Monterey
2. Chou PA (1991) Optimal partitioning for classification and regression trees. IEEE Trans Pattern Anal Mach Intell 13(4):340–354
3. Burshtein D, Pietra VD, Kanevsky D, Nadas AJ (1992) Minimum impurity partitions. Ann Stat 20(3):1637–1646
4. Nadas A, Nahamoo D, Picheny MA, Powell J (1991) An iterative flip-flop approximation of the most informative split in the construction of decision trees. In: International conference on acoustics, speech, and signal processing. Toronto, pp 565–568
5. Mehta M, Agrawal R, Rissanen J (1996) SLIQ: a fast scalable classifier for data mining. In: Proceedings of the fifth international conference on extending database technology. Springer, Berlin, pp 18–33
6. NASA (1992) Introduction to IND Version 2.1, GA23-2475-02 edition. NASA Ames Research Center

7. Coppersmith D, Hong SJ, Hosking JRM (1999) Partitioning nominal attributes in decision trees. Data Min Knowl Disc 3(2):197–217
8. Loh WY, Shih YS (1997) Split selection methods for classification trees. Statistica Sinica 7:815–840
9. Parsons l, Haque E, Liu H (2004) Subspace clustering for high dimensional data: a review. SIGKDD Explorations 6(1):90–105

Synthesized Algorithms of Concept Similarity Based on the Semantic Correlation Prerequisite

Hui-lin Liu and Qing Liu

Abstract This paper offers a synthesized approach of solving the shortage of the traditional similarity in ontology mapping. First, it selects high correlation concepts by Hirst-St-Onge semantic relativity algorithms, in order to reduce the complexity of the account. Then according to the characteristic of the ontology concept, we designs a synthesized method through calculating the respective similarity in name, attribute and instance of concepts, and works out weight by Sigmoid function. Experiment data indicates that it makes the better accuracy than the traditional methods.

Keywords Ontology mapping · Concept similarity · Semantic correlation · Weighting

1 Introduction

Since the 1990s, the research results of ontology have been widely used in artificial intelligence and semantic Web field, etc. Ontology, a method of conceptualizing and modeling domain knowledge, can make formalized description of knowledge understood by machine. It carries out the intelligent information interaction between machines. However, Web is decentralized when it refers to its framework. The developers in different areas or even in the same area use various ways to construct ontology, because the developers are diverse in language, culture and habit [1]. This

H. Liu (✉)
School of Computer Science and Engineering, Anhui University
and Technology, Huainan 232001, China
e-mail: liuhl@aust.edu.cn

Q. Liu
Jiangxi Application Engineering Vocational Institute, Pingxiang 337042, China

leads to the heterogeneity problem of ontology and betrays the original intention that the ontology is regarded as a means of knowledge reuse and share.

Ontology mapping means to set up mapping rules between heterogeneous ontology. Its essence is to search similar semantic concepts between two ontology. Therefore, we can think that the core work of ontology mapping is to compare the similarity degree of concepts between ontology. Nowadays, there are a lot of research literatures about calculation methods of concept similarity in ontology at home and abroad. But considering the engineering target of the application of ontology, we are in pursuit of methods with higher efficiency and better adaptability. Currently, there are mainly two problems in algorithm of concept similarity: First is that calculation is relatively large and of low efficiency, The second is that the calculation of the concept similarity is incomplete and the method of calculation is not perfect. Besides, sometimes we can't get accurate results because of the influence of people's subject factors.

In allusion to the above problems, this paper puts forward a new calculating method of concept similarity in ontology. After screening the concept in ontology which do not need to compare the similarity, we then calculate the similarity based on the element contained in the concept one by one. Afterwards, calculate the weight of similarity of each part by using the neural network function, and finally summate with weighting to obtain the final results.

2 Related Concepts

2.1 Ontology Definition and Composition

The exact definition about ontology, appear various from a slightly different version. But the most cited Studer [2] is summarized, such as the definition, namely "ontology is Shared conceptual model clear the formal specification", such as Perez with classification organized ontology, describe the ontology of five basic Modeling Primitive [3].

1. Classes or concepts: It refers to any affairs of abstract description, such as job descriptions, function, behavior, strategies and reasoning process, etc. The concept of ontology usually constitutes a category hierarchical relationship. In regard to Form, the concept is defined as a quad: C = {Id, L, P, I}. In this quard, Id is the only identifier of the concept which is described by URI; L is the language vocabulary of the concept, P is a set which has conceptual attribute; I belong to the collection of all instances of the concept.
2. Relation: That is the interaction between concepts in the fields.
3. Function: It is a kind of special relationship. The first element of this relation can only be determined by the former one.
4. Axiom: It represents some tautologies.
5. Instance: It represents specific entity referred by a concept.

2.2 Similarity of the Ontology Concept

When two ontological elements have some common features, they are defined as similar. The similar extent is described by similarity. In Form, the calculation of similarity meets: (1) $Sim(x,y) \in [0...1]$. The value of the similarity calculation is a real number of interval [0, 1]; (2) $Sim(x,y) = 1$, If and only if $x = y$, the two objects are completely similar, the similarity is 1; (3) $Sim(x,y) = 0$. If the two objects do not have any common feature, the similarity is 0; (4) $Sim(x,y) = Sim(y,x)$. The similarity relation is symmetrical.

3 Screening of Concept

Correlation refers to the relationship between concepts, which expresses the dependence between concepts. Correlation is more common than similarity. Similarity means the characteristics expressed by the concept overlap in some respects. Correlation indicates that there is some similarity between the concepts, but the expressions of characteristics do not necessarily overlap. As a result, similarity is a particular aspect of correlation. Resnik thinks that correlation with implication relationship equals to similarity. In normal conditions, similar concepts with similarity are regarded to have correlation, in other words, similar concepts are generally correlated, but correlative concepts are not necessarily similar, such as the keyboard and the mouse which have strong correlation but low similarity. This paper uses the Hirst-St-Onge semantic correlation algorithm [4], the basic idea is: when the two concepts with ontology synonym (Synset) have a shorter path to be connected, then we firmly believe that they have large semantic correlation relatively on semantics. Through the close relationship between the semantic similarity and semantic correlation, we can obtain that the two concepts are also likely to have greater semantic similarity. When this path does not exist, $Cr_{HS}(w_1, w_2) = 0$,

$$Cr_{HS}(w_1, w_2) = C - Len(w_1, w_2) - k \times d. \tag{1}$$

Among them, the $Cr_{HS}(w_1, w_2)$ is the definition of semantic correlation of Hirst-St-Onge, C and k are two constant parameters, d represents the steering number in Synset, $Len(w_1, w_2)$ represents the path length. The value range of $Cr_{HS}(w_1, w_2)$ is (0, 1). We'll give a constant r as threshold value, only when $Cr_{HS}(w_1, w_2) \geq r$, the two concepts are regarded to have concept correlation. Thus, we can determine that they may have concept similarity. Only when they satisfy the condition that $Cr_{HS}(w_1, w_2)$ can they proceed the concept similarity calculation.

4 Synthesized Calculation Method of Concept Similarity

According to the characteristic of the ontology concept, we designs a synthesized method through calculating the respective similarity in name, attribute and instance of concepts, and works out weight by Sigmoid function. At last, the final results of similarity of ontology concept are concluded. Among them, Sigmoid function produce the weight. Below introduces the implementation method of algorithm.

4.1 Similarity Algorithm of Concept Name

The theory of calculating the similarity through the concept name is based on: Normally if the name of the two concepts is represented by a same or similar character, then the meaning of the two concepts is usually same or similar. As a general rule, the similarity of the concept name uses "Edit Distance" algorithm to compare the string of the two concept names. Edit distance is measured by the least switching times needed to complete the conversion string. The main operation includes the inserting, deletion, replacement of the character as well as exchanging position with the adjacent character. The formula is:

$$\mathrm{Sim}(t_1, t_2) = \max(0, 1 - \frac{\mathrm{ed}(t_1, t_2)}{\min(|t_1|, |t2|)}. \tag{2}$$

Among them, t_1, t_2 are respectively the source string and target string to be compared, ed (t_1, t_2) represents editing distance of string t_1, t_2.$|t_1|$ is the number of characters in string t_1.

"Edit distance" is just a string matching method. Although this kind of practice is simple and clear, it lacks the consideration of entity of the semantic information. In reality, the homonymy and synonymy phenomenon is widespread. When the concept name of entity is just given by simple character instead of semantics, the matching method of calculating the similarity through characters is not reliable. Here, we adopt a method of combining common sense knowledge base of Hownet to carry out the semantic similarity calculation. Hownet is a common sense knowledge base, which is based on concepts that are represented by an English-chinese bilingual vocabulary and characteristics of the concepts. Its basic component is to reveal the relationship between concepts as well as between characteristics of concepts. Concept in Hownet is described by sememe, which is the minimum unit to describe concept. Suppose the path for two sememes in a tree hierarchy which is constituted by hyponymic relation is d, then the semantic similarity between the two sememes is:

$$\mathrm{Sim}(t_1, t_2) = \frac{\alpha}{d + \alpha}. \tag{3}$$

Among them, α is the sememe's semantic distance value when the semantic similarity is 0.5. But in Hownet, each concept is explained by more sememes, we have to calculate similarity of all the sememes in both concept A and B, and then do weighting to determine the similarity of concept name. Specific formulas are represented as below:

$$\mathrm{Sim}_{\mathrm{name}}(A, B) = \sum_{i=1}^{m} \beta_i \max_{1 \leq j \leq n} \mathrm{Sim}(t_i, t_j). \tag{4}$$

Among them, m, n is the sememe's number of A, B; β_i is the weighting $(\beta_1 + \beta_2 + \cdots + \beta_m = 1)$ of the ith sememe. It is given by the experts' experience.

4.2 Similarity Algorithm of Concept Attribute

According to the heuristic rules: if two concepts have the same attributes, these two concepts may be similar. According to attribute, we can calculate the similarity of concept. There are some elements in attribute like attribute name, attribute data type and attribute instance data. Whether two attributes is similar or not is mainly determined by the three elements above.

The front formula method can be used to calculate similarity degree of attribute name. In according to the data type matching table, we can find out matching degree K of two attributes a_i, a_j. Then we can obtain the combination similarity degree of attribute based on these two aspects:

$$\mathrm{Sim}(a_i, a_j) = \mathrm{Sim}_{\mathrm{name}}(a_i, a_j) \times K. \tag{5}$$

Suppose we calculate mth $\mathrm{Sim}(a_i, a_j)$ between concept A and concept B in total, and then set up corresponding weights, thus the similarity degree based on attribute between concepts is:

$$\mathrm{Sim}_{\mathrm{attribute}}(A, B) = \frac{\sum_{g=1}^{m} \omega_g \mathrm{Sim}(a_i, a_j)}{\sum_{g=1}^{m} \omega_g} \tag{6}$$

Here, we need to consider that because a concept may have multiple attributes, the degree and function of concept's description made by each attribute is different from each other. If we take every attribute into consideration, the calculation will be very large. So while in the calculation of attribute's similarity, we should determine the priority of each attribute first. This paper uses inductive method of Decision Tree in data mining, in line with a set of concepts instances to calculate each attribute's information gain. We are about to rank the attributes according to the decreasing of information gain [5], only selecting several attributes of large

information gain to calculate similarity. It can reduce the calculation amount in a certain extent. The definition and calculating method of information gain is as follows: If S is a set of training sample, S contains S_ith C_i sample, and attribute A can be used to divide S into subset {s1, s2,..., sn}, The information gain of attribute A received by this division is defined as: Gain $(A) = I(s1,s2,...,sn) - E(A)$. Among them, Gain(A) = I(s1, s2,..., sn) is the expectation for classification of a given sample, E(A) is the entropy of attribute A.

4.3 Similarity Algorithm of Concept Instance

We can use concrete instance to calculate concept similarity of two ontology which are in need of ontology mapping. Its theoretical basis is the heuristic rules: If two concepts have the same instance, they may be similar; Or if the two concepts have a large proportion of the same instance, thus these two concepts may be similar. We use Jaccard-coefficient to calculate the similarity between concept instance, and the formula is:

$$\text{Sim}_{\text{instance}}(A,B) = \frac{P(A \cap B)}{P(A \cup B)} = \frac{P(A,B)}{P(A,\overline{B}) + P(A,B) + P(\overline{A},B)} \tag{7}$$

Use positive and negative sample combined with machine learning methods to train learning devices, then calculate the value of $P(A,B), P(A,\overline{B})$ and $P(\overline{A},B)$ in formula.

4.4 The Weighted Synthesis and Weight Calculation of Similarity

In order to get a comprehensive calculation results, it is necessary to combine several measurement results. There exists much combination method, and in this paper, we use composite which is a very popular method in current. The specific calculation formula is listed below:

$$\text{Sim}_{\text{composite}}(A,B) = \sum_{i=1}^{3} \varphi_i \cdot X_i \tag{8}$$

Among them, $X_1 = \text{Sim}_{\text{name}}(A,B)$, $X_2 = \text{Sim}_{\text{attribute}}(A,B)$, $X_3 = \text{Sim}_{\text{instance}}(A,B)$. φ_i is the weight, which is produced by sigmoid function. Sigmoid function is a kind of smoothing effect function, which makes the combination result in favour of a strategy with high predicted value. Here we adopt this form:

$$\text{Sig}(X_i) = \frac{1}{1 - e^{-5(x-0.5)}}. \tag{9}$$

The weight φ_i is defined as follows:

$$\varphi_i = \frac{\text{Sig}(X_i)}{\sum\limits_{i=1}^{3} \text{Sig}(X_i)}. \tag{10}$$

5 Experimental Analysis

5.1 The Experiment Data and Platform

The experiment in this paper is aimed at tourist resources ontology established by the research team of tsinghua university. The ontology not only defines the concepts and hierarchical relationships of food, housing, transportation, travel, purchasing, etc., related to travel, but also makes use of the ontology development tool Protégé to create ontology (concept) instances.

This algorithm is carried out by using VC++ language to write code, and runs on the Winows XP.

5.2 The Experimental Results and Evaluation

Here, we use this algorithm (Algorithm 2) to compare with algorithm (Algorithm 1) which is proposed by Qun and Sujian [6]. Besides, algorithm 1 is based on sememe to calculate concept similarity in Hownet (While in calculation, in algorithm 2, the threshold value of correlation is set as 0.3. If the concept interval is below this value, the similarities are all set as 0.1. The rule of the parameters used in algorithm 1 are set as follows: $\alpha = 1.8$, $\beta_1 = 0.5$, $\beta_2 = 0.2$, $\beta_3 = 0.17$, $\beta_3 = 0.17$, $\beta_4 = 0.13$, $\lambda = 0.2$.), The experimental results are in the Table 1:

Table 1 Comparison of experimental results of similarity

Concept 1	Concept 2	Algorithm 1	Algorithm 2	
			Correlation	Similarity
Currency	Credit card	0.4444	0.5670	0.4633
Airplane	AirPort	0.2424	0.8127	0.2790
Ticket	Postcode	0.1412	0.1750	0.1000
Hotel	Restaurant	1.0000	0.7250	0.9685
Car	Gas station	0.3957	0.7890	0.2310

It can be seen from the experimental results that the algorithm 1 is rough in measuring similarity. For example, the correlation of "Car" and "Gas station" is very high. But when judged by experience in this field, the similarity degree is not high. What's more, from the point of view of concept instance, "Hotel" and "Restaurant" are not completely similar. The improved algorithm put forward by this paper is relatively reasonable.

6 Conclusion

This paper puts forward a new comprehensive calculation method of concept similarity. This method determines the correlation of concept, filtering out a large amount of irrelevant concepts. Thus it can reduce quite a lot unnecessary calculation; In the calculation of concept similarity, it takes influence of concept instances on concept similarity into fully consideration. This action makes the calculating results reflect the similarity relationship between concepts in a more comprehensive and accurate way. But the method in this paper need to be further developed in some place, for example, it is far from enough to determine the relationship between correlation and similarity only through one threshold value. Therefore, the next step is to continue improving the judgment rules of concept correlation.

References

1. Jian-hou G, Yue J, You-Ming X (2011) Ontology and its application. Science Press, Bei Jing
2. Studer R, Benjamins VR, Fensel D (1998) Knowledge engineering: principles and methods. Data Knowl Eng 25(122):161–197
3. Zhong-ping Z, Shu-xia T, Hong-qiang L (2008) Compositive approach for ontology similarity computation. Comput Sci 35(12):142–182
4. Song W, Yong M, Gang W (2010) An efficient heterogeneous ontology mapping algorithm. Comput Eng Sci 32(10):93–96
5. Zhi-feng G, Yong L, Gen-cheng G (2008) Improved method of conceptual similarity in ontology mapping. Comput Eng Appl 44(8):67–70
6. Qun L, Su-jian L (2002) Caculation of Semantic Similarity of Vocabulary Based on "Hownet". Comput Linguist Chin Lang Process 7(2):59–76

A Framework for Density Weighted Kernel Fuzzy *c*-Means on Gene Expression Data

Yu Wang, Maia Angelova and Yang Zhang

Abstract Clustering techniques have been widely used for gene expression data analysis. However, noise, high dimension and redundancies are serious issues, making the traditional clustering algorithms sensitive to the choice of parameters and initialization. Therefore, the results lack stability and reliability. In this paper, we propose a novel clustering method, which utilizes the density information in the feature space. A cluster center initialization method is also presented which can highly improve the clustering accuracy. Finally, we give an investigation to the parameters selection in Gaussian kernel. Experiments show that our proposed method has better performance than the traditional ones.

Keywords Fuzzy clustering · Parzen density estimation · Gene expression data

1 Introduction

Clustering has been used in gene expression data analysis for many decades, which is an unsupervised learning technique [1]. It not only identifies biologically relevant of genes, but also predicts the unknown gene's function. A wide range of clustering algorithms have been proposed [1–6]. The most widely used clustering techniques are partitional clustering such as *k*-means [2], self-organizing map [3], and hierarchical clustering [4]. However, both of them suffer many limitations. The parrtitional clustering is based on the pair-wise distance; the choice of

Y. Wang (✉) · M. Angelova
Mathematical and Intelligent Modeling Lab, Northumbria University,
Newcastle upon Tyne, Tyne and Wear, UK
e-mail: yu.wang@unn.ac.uk

Y. Zhang
Computational Intelligent Lab, Northumbria University, Newcastle upon Tyne,
Tyne and Wear, UK

distance similarity plays a key role in clustering efficiency. When the clusters differ in size and shape, most of the patitinal clustering algorithms will lose efficiency. The hierarchical clustering is also problematic due to it usually generates one single large cluster and several singletons. Moreover, all of these methods assign genes to exactly one cluster even if their expression profile is similar to several cluster patterns [5]. Fuzzy c-means (FCM) was proposed which assign each data a membership degree to a cluster. The membership values vary between zero and one. This feature enables fuzzy clustering are more appropriate for gene expression data. However, fuzzy c-means still suffers some weaknesses. First, gene expression data includes large proportion of noise, FCM neglects this characteristics and assigns equal weights to the meaningful data and the noise. This simple treatment leads to an inaccurate result. In addition, FCM is only effective to find the crisp, spherical clusters, but gene expression data is complicated which includes many non-spherical and overlapping structure. This complexity will reduce the effectiveness of FCM largely.

In this paper, we propose a clustering technique based on the original fuzzy c-means. In order to reduce the noise influence in gene expression data and increase the robustness, we utilize density information in the kernel space by adding a weighted parameter to discern the noise and meaningful data. In addition, we propose an initial cluster center selection to reduce the initialization sensitivity. Finally, we give an investigation to the parameters in Gaussian kernel. Experiment results show our proposed method outperform than other conventional methods.

The rest of the paper is organized as follows, in Sect. 2, we give an overview of the original fuzzy clustering, kernel fuzzy clustering. Section 3 considered the new proposed clustering method (density weighted kernel fuzzy c-means) and the initial cluster center selection. Validation measures and experiments results are given in Sect. 4. Finally, conclusion is presented in Sect. 5.

2 Backgrounds

2.1 Fuzzy c-Means

The fuzzy c-means is an extension of hard k-means clustering, in which a sample can be assigned to more than one clusters with a membership degree [5]. Given a data set $X = \{x_1, x_2, \ldots, x_n\} \subset R^{N \times Q}$, N is the number of samples, Q is the dimension of the sample x_j. The fuzzy c-means is an iterative optimization that minimizes the cost function defined as follows:

$$J = \sum_{i=1}^{C} \sum_{j=1}^{N} u_{ij}^m \|x_j - v_i\|^2 \tag{1}$$

where C denote the number of cluster, u_{ij} represent the membership of point x_j in the ith cluster, satisfying $0 \leq u_{ij} \leq 1$ and $\sum_{i=1}^{c} u_{ij} = 1$. v_i is the ith cluster center. $\|.\|$ is a norm metric. The parameter m controls the fuzziness of the resulting partition. In this work, we set $m = 2$.

The traditional FCM uses the squared norm to measure the similarity between centers and samples, which is effective in finding spherical clusters. However, the relations between genes are often non-linear and many clusters are non-spherical and overlapped, so the performance of this choice is not satisfactory. In order to cluster more general data set, many methods have been proposed by replacing the squared-norm in Eq. 1 with other similarity measures, such as *Pearson correlation*, *Mutual information* etc. A novel method is to use the kernel function to construct a kernel version of fuzzy clustering algorithm.

2.2 Kernel Fuzzy c-Means

Kernel fuzzy c-means is a novel clustering algorithm, which maps the samples to a higher dimensional feature space using a non-linear function, then performs fuzzy c-means in the feature space. If the nonlinear function is smooth and continuous, then the characteristic and the structure of the data in the original space will be preserved in the feature space [6].

Define a nonlinear map $\phi : x \rightarrow \phi(x) \in F$, where $x \in X$, X and F denotes the data space and the feature space respectively. F is usually a higher or even infinite dimension. The kernel fuzzy clustering method minimizes the following objective function

$$J = \sum_{i=1}^{C} \sum_{j=1}^{N} \mu_{ij}^{m} \left\| \phi(x_j) - \phi(v_i) \right\|^2 \tag{2}$$

3 Proposed Method

3.1 Initialization of Cluster Centers

The performance of FCM and KFCM algorithms highly depend on the selection of initial cluster centers. The conventional method is to run the algorithm several times with different initialization. However, this approach has the drawback of being time consuming and computationally expensive.

As the high density area contains more samples, "good" initial centers usually exist in these areas. We should select the samples with high density values as the initial cluster centers. The density value can be estimated by the Parzen density

function. The selection of initial cluster centers is basically a selection of C local extreme values of the sample's density.

Parzen density function is a nonparametric estimation technique which can estimate the sample's density without any prior assumption on its form. Given a set of N samples $X = \{x_1, x_2,\ldots, x_N\}$, the cluster number is c. To estimate the density of x is

$$p(x) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{V}\varphi\left(\frac{x - x_i}{h}\right) \tag{3}$$

where $p(x)$ is density estimation of x, N is the number of samples, V is the volume of the hypercube centering at x. h is the radius. $V = h^d$. $\varphi(u)$ is the window. The frequently used windows are rectangular, normal and exponential windows.

The parameter u in the Parzen function can be calculated by Eq. 4:

$$u = \frac{x - x_i}{h} \tag{4}$$

when x_i fall into the hypercube which centers at x, the radius is h, then $\varphi(u) = 0$, else $\varphi(u) = 0$. Therefore, the number of the samples in the hypercube is:

$$k = \sum_{i=1}^{N}\varphi\left(\frac{x - x_i}{h}\right) \tag{5}$$

Compared to the Euclidean distance, kernel distance is more robust to the noise, which is more appropriate for gene expression data. Therefore, the selection of initial centers should be calculated in the kernel space. Define ϕ is the kernel function, Eq. 4 should be revised to:

$$u = \frac{\phi(x) - \phi(x_i)}{h} \tag{6}$$

$$\|u\|^2 = \frac{1}{h^2}(\phi(x) - \phi(x_i))^T(\phi(x) - \phi(x_i)) = \frac{1}{h^2}(k(x,x) - 2k(x,x_i) + k(x_i,x_i)) \tag{7}$$

If we use Gaussian kernel, $k(x, x) = 1$, the Eq. 7 is simplify:

$$\|u\|^2 = \frac{2}{h^2}(1 - k(x,x_i)) \tag{8}$$

The density estimation of each x_i in the Gaussian kernel space is:

$$p(x_i) = \frac{1}{NV\sqrt{2\pi}}\sum_{j=1}^{N}\left(\exp\left(-\frac{1}{h^2}(1 - k(x_j,x_i))\right)\right) \tag{9}$$

The density estimation of each data point actually can be seen as the value of potential function [7]. Define potential function of x_i:

$$P_i^{(0)} = \frac{1}{NV} \sum_j^N \varphi(u) \tag{10}$$

Define $P_1^* = \max\left\{P_i^{(0)}, i = 1, 2, \ldots, N\right\}$, take x_1^* as the first cluster center, then adjust the potential value of each point according to Eq. 11:

$$P_i^k = P_i^{k-1} - P_k^* \frac{1}{1 + \left\|x_i - x_k^*\right\|^2}, \quad k = 1, 2, \ldots, c - 1 \tag{11}$$

3.2 Density Weighted Kernel Fuzzy c-Means

The objective function of density weighted kernel fuzzy c-means is defined as follows:

$$J = \sum_{i=1}^{C} \sum_{j=1}^{N} w_j \mu_{ij}^m \left\|\phi(x_j) - \phi(v_i)\right\|^2 \tag{12}$$

A new set of parameters w_j are introduced to the objective function, which is based on the characteristic and structure of data set. w_j is calculated by the normalization of the density information,

$$w_j = \frac{p(x_j)}{\sum_{j=1}^{N} p(x_j)} \tag{13}$$

By utilizing Lagrange multipliers, the optimal fuzzy partition matrix U and the optimal cluster center matrix V are obtained in Eqs. 14 and 15 respectively,

$$u_{ij} = \frac{(1/(1 - K(x_j, v_i)))^{1/(m-1)}}{\sum_{i=1}^{c} (1/(1 - K(x_j, v_i)))^{1/(m-1)}} \tag{14}$$

$$v_i = \frac{\sum_{j=1}^{n} w_j u_{ij}^m K(x_j, v_i) x_j}{\sum_{j=1}^{n} w_j u_{ij}^m K(x_j, v_i)} \tag{15}$$

3.3 Parameter σ Selection

An optimal selection of σ can improve the clustering accuracy substantially. Most investigation on the σ are based on SVM, however, the parameter selection based on FCM rarely can be seen. Li [8] proposed an auto parameter selection on SVM based on the training sets. Using his idea, we select k nearest samples around the

initial cluster center to form a training set X', the cluster information in X' is obvious known. According to the attributes of the Gaussian kernel function,any two data points in the data set X' should satisfy Eq. 16:

$$\begin{cases} k(x,z) \approx 1, x,z \in \omega_i \\ k(x,z) \approx 0, x \in \omega_i, z \in \omega_i, i \neq j \end{cases} \tag{16}$$

where ω_i denotes ith cluster. Li [8] also gave two functions on σ, which is an independent variable:

$$w(\sigma) = \frac{1}{\sum\limits_{i=1}^{c} N_i^2} \sum_{i=1}^{c} \sum_{x \in \omega_i} \sum_{z \in \omega_i} k(x,z,\sigma) \tag{17}$$

$$b(\sigma) = \frac{1}{\sum\limits_{i=1}^{c} \sum\limits_{\substack{j=1 \\ i \neq j}}^{c} N_i N_j} \sum_{i=1}^{c} \sum_{\substack{j=1 \\ i \neq j}}^{c} \sum_{x \in \omega_i} \sum_{z \in \omega_j} k(x,z,\sigma) \tag{18}$$

The selection of σ is formulated as a minimization of $J(\sigma)$

$$\min_{\sigma > 0} J(\sigma) = 1 - w(\sigma) + b(\sigma) \tag{19}$$

4 Data and Experiment Results

4.1 Data Sets

Yeast cell cycle data: we use a subset of *Yeast cell cycle data*, which contains 237 genes corresponding to four categories in the MIPS database [9].

   Rat CNS: This data is to study the gene expression level during rat central nervous system development [10], which contains 112 genes over 17 conditions.

4.2 Evaluation Measure

External validation: Adjusted rand index is a measure of the agreement between two partitions [11]. It varies from -1 to 1 and higher value indicates that the clustering result is more similar to the standard partitions. In addition, we compare the cluster's biological functional consistency by David [12], which is an online analysis tool to provide cluster's functional interpretation using Gene Ontology (*GO*).
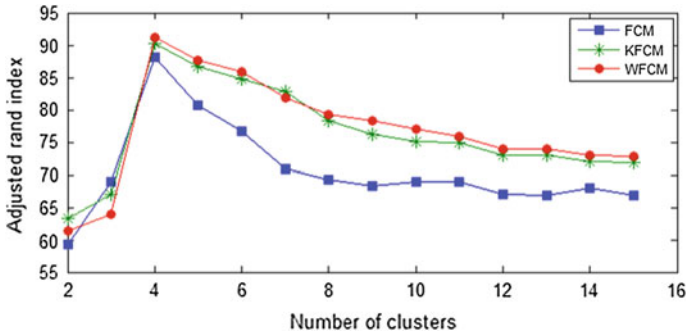
Fig. 1 Adjusted rand index for *Yeast 237*

4.3 Clustering Results

1. Adjusted rand index

Figure 1 shows the adjust rand index over various number of clusters in *Yeast 237*. It can be seen that all of the three methods can detect the optimal number of clusters. Moreover, the clustering result by our proposed method is more close to the standard partition. Similar results can be found in Fig. 2 based on *rat CNS*.

2. Biological consistency of the clustering result

We calculate annotation ratio for each cluster and use the average value as the final ratio. The annotation ratio is based on three structured vocabularies: biological processes (*BP*), cellular components (*CC*) and molecular functions (MF). A high value indicates that the corresponding clustering method is better correlated with biological functions. Figures 3 and 4 illustrates the three methods annotation ratios on *Yeast 237 and rat CNS* respectively. The annotation ratios generated by our proposed method (WFKCM) outperforms the other methods.
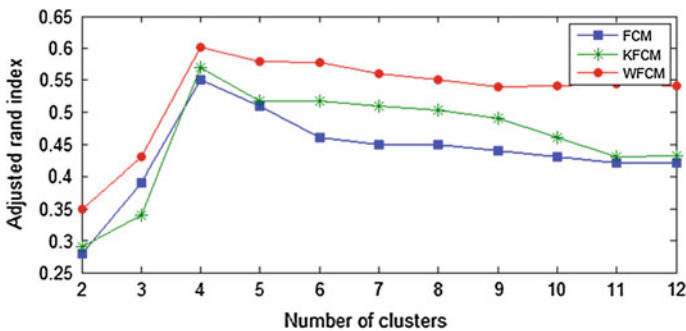
Fig. 2 Adjusted rand index for *rat CNS*

Fig. 3 Biological annotation ratios on *Yeast 237*
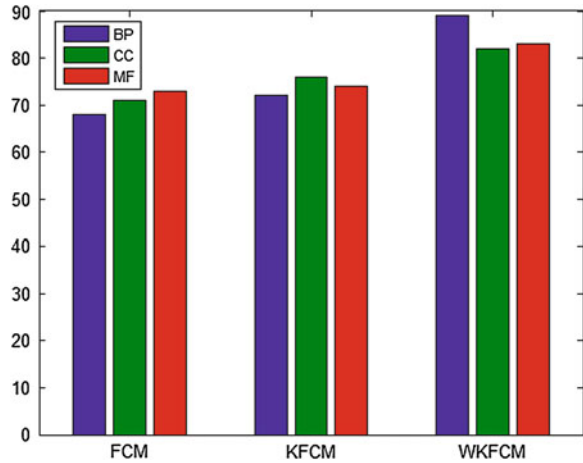
Fig. 4 Biological annotation ratios on *rat CNS*
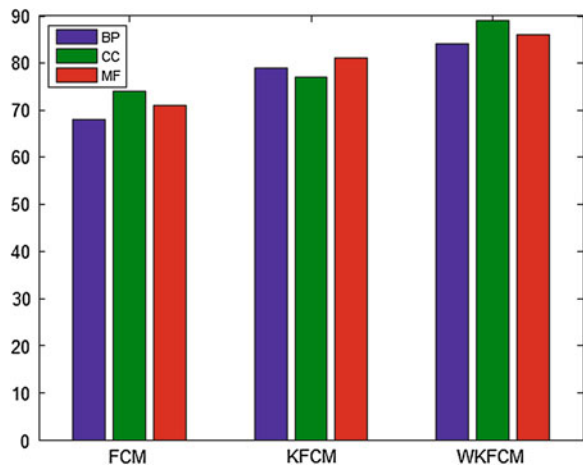
5 Conclusion

In this paper, we have proposed a new weighted fuzzy clustering method for gene expression analysis, which employed density information in kernel space to reduce the influence of noise. In addition, an initial cluster center selection was presented which utilize the potential function by searching local extreme density values. Finally, we gave an investigation to the parameter σ in Gaussian kernel and proposed an optimization method. Experiment results show that our proposed method has better performance than the conventional FCM and KFCM on gene expression data.

References

1. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95:14863–14868
2. Hartigan JA, Wong MA (1979) A *k*-means clustering algorithm. Appl Stat 28:100–108
3. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display genome-wide expression patterns. In Proc Natl Acad Sci, pp 14863–14868
4. Guess MJ, Wilson SB (2002) Introduction to hierarchical clustering. J ClinNeurophysiol 19:144–151
5. Dembele D, Kastner P (2003) Fuzzy *c*-means method for clustering microarray data. Bioinformatics 19(8):973–980
6. Camastra F, Verri A (2005) A novel kernel method for clustering, IEEE Trans Pattern Anal Mach Intell 27(5):801–805
7. Chiu SL (1994) Fuzzy model identification based on cluster estimation. J Intel Fuzzy Syst 2(3):267–278
8. Li CH, Lin CT, Kuo BC, Chu HS (2010) An automatic method for selecting the parameter of the RBF kernel function to support vector machines. IEEE IIGARSS, pp 836–839
9. Tavazoie S et al (1999) Systematic determination of genetic network architecture. Nat Genet 22:281–285
10. Wen X, Fuhrman S, Michaels GS (1998) Large-scale temporal gene expression mapping of central nervous system development. Proc Natl Acad Sci 95:334–339
11. Yeung KY, Haynor DR, Ruzzo WL (2001) Validating clustering for gene expression data. Bioinformatics 32:131–138
12. Dennis J, Sherman BT, Hosack DA, Yang J (2003) DAVID: database for annotation, visualization, and integrated discovery. Genome Biol 4(5):3–9

Simulation of Extracting Phase from the Interferogram by the Phase Shifting Method

Zhanrong Zhou, Aijun Li, Hongxia Wang, Yunfang Zhao and Jin Ma

Abstract The phase information extraction from the interference fringe is a key step in optical interferometry. Phase shift extraction method is proposed. This method is introduced amount of phase shift in the reference light, changing the relative phase of two coherent wavefront, comparing with a different amount of phase shift value in the intensity light to solve the points phase. The four steps phase shifting extraction is simulated. The simulation results show that the method not only has a strong ability of the inhibition noise, but also can identify the arbitrarily shaped interference fringes.

Keywords Information optics · Computer simulation · Interference fringes · Phase extraction

1 Introduction

In optical interferometry, the information of physical quantity to be measured implicit in the phase of the light waves, but measured results are recorded in the form of interference fringes. How to extract the phase information from the interference fringe is a key step to obtain the measurement results. Therefore the research of extracting phase information from the interference fringes has become one of the main topics in optical interferometry [1–5]. The method of extracting phase from interference fringes is mainly divided into two categories [6, 7]: One is Fourier transform method from a single fringe image. Firstly it is transform the interference fringes by Fourier transformation, secondly take out a spectrum, finally inverse Fourier transform to extract the phase distribution. Although it need only one fringe

Z. Zhou (✉) · A. Li · H. Wang · Y. Zhao · J. Ma
Department of Physics Xi'an High Technology Studio, Xi'an, 710025, China
e-mail: zhouzhou76@163.com

image, it are more difficult to determine a spectrum and band width of the filter, and can only analyze the monotonous fringe image, which makes Fourier transform method's application range limited. The other is the phase shift method from many fringe images. It is introduced amount of phase shift in the reference light, changing the relative phase of two coherent wavefront, comparing with a different amount of phase shift value in the intensity light to solve the points phase in the interference field [8, 9].

The theory of the phase shift is analyzed in detail in this paper. This method not only can analyze non monotonic fringe image, but also has the ability to recognize increase stripes. Phase extraction is simulated on the basis of phase shift, which is the foundation of the actual optical interferometry.

2 Basic Principle

The phase shift method to extract the phase comes from the mathematical description of interference fringe, for the vast majority of optical interferometry, the image of interference fringes can be expressed as

$$I(x, y) = a(x, y) + b(x, y) \cos[\phi(x, y)] \tag{1}$$

In the formula (1), $a(x, y)$ is the background, $b(x, y)$ is a light wave amplitude, $\phi(x, y)$ is the wave phase.

Due to the typically stripe represents the physical quantity to be measured can be expressed as a phase function of $\phi(x, y)$, but not to rely on $a(x, y)$ and $b(x, y)$. So if we can get $\phi(x, y)$, we can obtain the physical quantity to be measured results. Therefore the interference fringe phase extraction is the use of the stripe image gray level information to obtain fringe phase distribution. If the phase distribution is calculate from a fringe pattern. It will be faced with three unknown difficulties from an equation.

Through the interference method to measure phase, it must be added known conditions on the basis of Eq. (1). In the practical measuring light path, the phase shifter is used to increase the number of phase in order to get more interference fringes, which to solve the phase distribution [10].

If the corresponding phase displacement φ_i is accurately known, the tested wave-front phase distribution is Φ, the phase distribution $\phi(x, y)$ in light field can be expressed as $\phi(x, y) = \Phi + \varphi_i$, the exponential form of formula (1) is expressed into

$$I_i = a + \frac{b}{2} \exp[-j(\Phi + \varphi_i)] + \frac{b}{2} \exp[j(\Phi + \varphi_i)] \tag{2}$$

The intensity sample function in reference phase range can be expressed as

$$I_s = \sum_{i=1}^{N} I_i \delta(\varphi - \varphi_i) \tag{3}$$

The Fourier transform of formula (3) is

$$\tilde{I}_s = \int_0^{2\pi} I_s(\varphi) \exp[-ju\varphi] d\varphi = \sum_{i=1}^{N} I_i \exp[-ju\varphi_i] \tag{4}$$

Formula (4) into the formula (3) can be

$$\tilde{I}_s(u) = A \sum_{i=1}^{N} \exp[-ju\varphi_i] + B \sum_{i=1}^{N} \exp[-j(u+1)\varphi_i] + C \sum_{i=1}^{N} \exp[-j(u-1)\varphi_i] \tag{5}$$

In the formula, $B = \frac{b}{2}\exp[-j\Phi], C = \frac{b}{2}\exp[j\Phi]$
Comparison of formula (4) and (5) can be

$$\sum_{i=1}^{N} I_i \exp[-ju\varphi_i] = A \sum_{i=1}^{N} \exp[-ju\varphi_i] + B \sum_{i=1}^{N} \exp[-j(u+1)\varphi_i]$$
$$+ C \sum_{i=1}^{N} \exp[-j(u-1)\varphi_i] \tag{6}$$

In the formula, I_i is the measured light intensity value, A, B, C are three unknowns
If selected the appropriate three frequencies u_1, u_2, u_3, A, B, C are solvable. i.e.

$$\begin{bmatrix} A \\ B \\ C \end{bmatrix} = H^{-1} \begin{bmatrix} \frac{1}{N}\sum_{i=1}^{N} I_i \exp[-ju_1\varphi_i] \\ \frac{1}{N}\sum_{i=1}^{N} I_i \exp[-ju_2\varphi_i] \\ \frac{1}{N}\sum_{i=1}^{N} I_i \exp[-ju_3\varphi_i] \end{bmatrix} \tag{7}$$

In formula (7)

$$H = \begin{bmatrix} \frac{1}{N}\sum_{i=1}^{N}\exp[-ju_1\varphi_i] & \frac{1}{N}\sum_{i=1}^{N}\exp[-j(u_1+1)\varphi_i] & \frac{1}{N}\sum_{i=1}^{N}\exp[-j(u_1-1)\varphi_i] \\ \frac{1}{N}\sum_{i=1}^{N}\exp[-ju_2\varphi_i] & \frac{1}{N}\sum_{i=1}^{N}\exp[-j(u_2+1)\varphi_i] & \frac{1}{N}\sum_{i=1}^{N}\exp[-j(u_2-1)\varphi_i] \\ \frac{1}{N}\sum_{i=1}^{N}\exp[-ju_3\varphi_i] & \frac{1}{N}\sum_{i=1}^{N}\exp[-j(u_3+1)\varphi_i] & \frac{1}{N}\sum_{i=1}^{N}\exp[-j(u_3-1)\varphi_i] \end{bmatrix} \tag{8}$$

By the formula (7) and (8), B and C can be calculated, and then the measured phase distribution can be calculated [11, 12].

$$\Phi = -\arctan\left[\frac{\mathrm{Im}(B)}{\mathrm{Re}(B)}\right] \tag{9}$$

$$\Phi = -\arctan\left[\frac{\mathrm{Im}(C)}{\mathrm{Re}(C)}\right] \tag{10}$$

$$\Phi = -\arctan\left[\frac{\mathrm{Im}(C/B)}{\mathrm{Re}(C/B)}\right] \tag{11}$$

3 Algorithm to Extract Shift Phase

Usually the phase shifter uses interval step sampling, reference phase φ_i is expressed as:

$$\varphi_i = 2\pi(i-1)/N \quad (i = 1, 2, \cdots, N) \tag{12}$$

The matrix H in Formula (8) can be converted into unit matrix:

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{13}$$

It calculates the phase for:

$$\Phi = -\arctan\frac{\sum\limits_{i=1}^{N} I_i \sin\left[\frac{2\pi(i-1)}{N}\right]}{\sum\limits_{i=1}^{N} I_i \cos\left[\frac{2\pi(i-1)}{N}\right]} \tag{14}$$

Equation (15) is the basic equation of phase shift interferometry.

When N = 3, the amount of phase shift can be respectively 0, $\frac{2\pi}{3}$, $\frac{4\pi}{3}$, and then the phase shifting interferometry basic equation can be simplified to

$$\Phi = -\arctan\frac{\sqrt{3}I_2 - \sqrt{3}I_3}{2I_1 - I_2 - I_3} \tag{15}$$

When N = 4, the amount of phase shift can be respectively 0, $\frac{\pi}{2}$, π, $\frac{3\pi}{2}$, and then the phase shifting interferometry basic equation can be simplified to

$$\Phi = -\arctan\frac{I_2 - I_4}{I_1 - I_3} \tag{16}$$
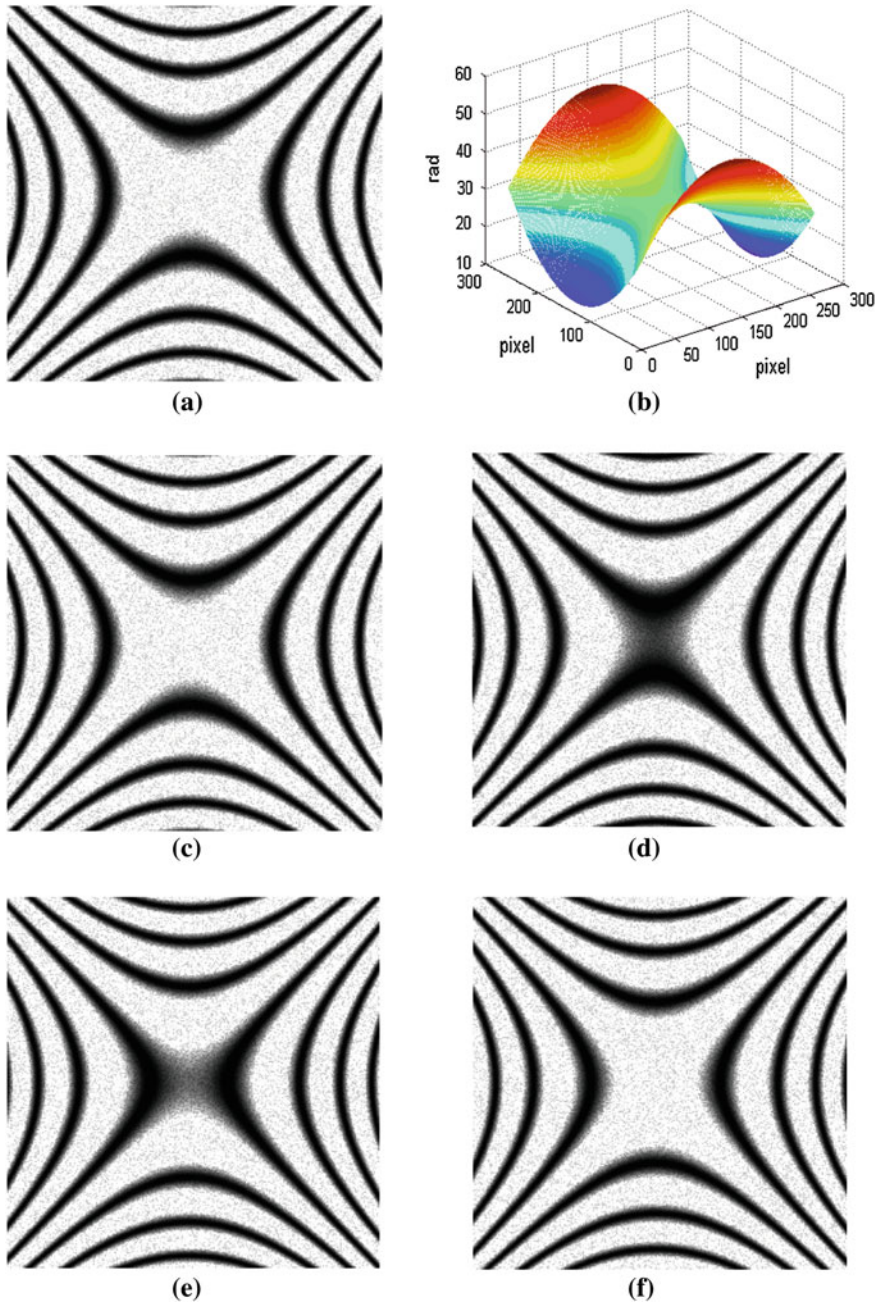
Fig. 1 4-steps phase shifting (**a**) simulation of fringe with noises (**b**) the distribution of fringe phase (**c**) phase shifting 0 (**d**) phase shifting (**e**) phase shifting π (**f**) phase shifting $3\pi/2$

Based on the analysis of the basic equation of phase shift interferometry, phase shift algorithm has the following advantages:

1. as a result of using subtraction, phase shift algorithm is not sensitive to noise, and requires lowly contrast and uniformity;
2. Because phase shift algorithm use point-to-point operations in phase processing, it can process arbitrary fringe image, regardless of whether the monotonic changes in fringe phase;
3. Because phase shift algorithm is taken directly measuring the fringe phase, its accuracy can reach one of tens to hundreds of a stripe cycle.

4 Verification

In order to verify the phase shift algorithm for noise suppression ability and stripe shape requirements, the phase of non monotone fringe image with noise is simulated. The simulation results is shown in Fig. 1: (a) is a simulation of the phase non monotone interference fringe with noise. (b) is the phase distribution of simulated fringe, the phase shift amount $(0, \pi/2, \pi, 3\pi/2)$ are applied in the reference light. To obtain the corresponding image of interference fringes are graph (c), (d), (e), (f).

The wrapped phase distribution is obtained using of formula (16), and shown in Fig. 2. The phase value is wrapped in the main value range of the inverse trigonometric function by the formula (16). Its value is truncated in $[-\pi, +\pi]$, but the actual measured phase change is often beyond the range. In order to obtain the true phase distribution, it is need that the wrapped phase is unwrapped. The wrapped phase is unwrapped by the ranks of point-by-point algorithms in this paper. Phase

(a)   (b)
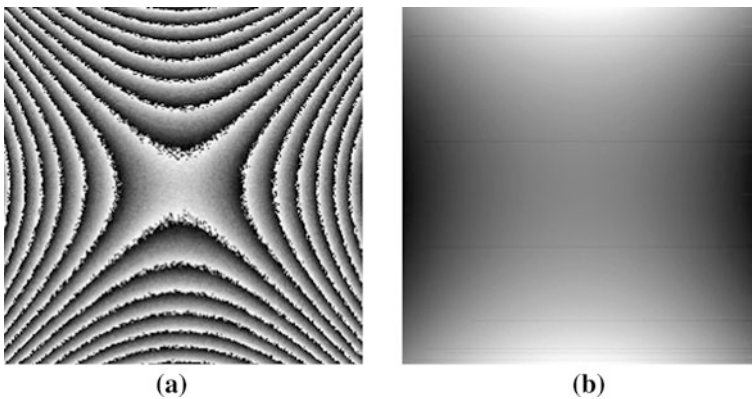
Fig. 2 Wrapping phase and unwrapping phase (**a**) the wrapping phase distribution (**b**) the unwrapping phase distribution

unwrapping results is shown in Fig. 2b. From the results of phase unwrapping, there is almost without noise in the extracting wrapped phase by the phase shift algorithm.

5 Conclusion

The phase shift algorithm is analyzed in detail and is simulated. The results show that the phase shift method not only has a strong ability of the inhibition noise, but also has the ability to identify the arbitrarily shaped interference fringes. The studies provide the foundation of the practical optical interferometry.

References

1. Li J, Su XY, Guo LR (1990) An improved Fourier transform profilometry for the automatic measurement of 3D object shapes. Opt Eng 29:1439–1444
2. Lai G, Yatagai T (1991) Generalized phase-shifting interferometry. Opt Soc Am, A8:822–827
3. Zhang Q, Xian-feng XU, Hong-guang Y, Lu Guang-can, Jiao ZY (2011) Phase-shift extraction and wave reconstruction in Four-step phase-shifting interferometry. Opo-Electronic Eng, 38(8):139–144
4. Mahapatra Sabyasachi (2012) Phase shift between supersymmetric partner potentials. Sci Res Pub J 3:74–79
5. Reid GT (1987) Automatic fringe pattern analysis: a review. Opt Lasers Eng 7:37–68
6. Xian-feng X, Li-li H, Hong-guang Y (2011) Analysis and correction of object wavefront reconstruction errors in two-step phase-shifting interferometry. Acta Phys Sin 60(8):084206
7. Cai CQ, He LF (2011) Phase difference extraction based on Four-step phase shifting. J South Chin Univ Technol 39(9):93–96
8. Schwider J (1993) New compensating four-phase algorithm for phase-shift interferometer. OE 32:1883–1885
9. Kemao Q, Fangjun S, Xiaoping W (2000) Determination of the best phase step of Carre algorithm in phase shifting interferometry. Meas Sci Tech 8:1120–1123
10. Zhang Y, Wang D, Zhao J (2009) Investigation on phase unwrapping algorithms in digital holography. ACTA Optica sinica 29:3323–3327
11. Xiaofan Q, Zhang Y, Li X et al (2010) Phase unwarpping algorithm based on mask and least squares iteration. ACTA Optica sinica 30:441–445
12. Takeda M, Ina H, Kobayashi S (1982) Fourier-transform method of fringe pattern analysis for computer-based topography and interferometry. J Opt Soc Am, 2:156–160

Research and Implementation of CAPTCHA Based on Ajax

Xiaokui Chen and Xinqing Lin

Abstract Currently, most authentication system requires users to answer the CAPTCHA (Completely Automated Public Turing Test to Tell Computer and Human Apart) before gaining the system access. CAPTCHA is a standard security technology for distinguish between human and computer program automatically. This paper describes the Ajax technology and its advantages of web interactivity and efficiency of the user to enter the CAPTCHA may appear misjudgment the inconvenience. Designed Ajax-based CAPTCHA scheme which can prompt the user input of right or wrong. This scheme using the characteristics of Ajax technology asynchronous transmission, after the user input CAPTCHA, immediately the prompt of entry correct or not gave. Greatly increase the website interactivity, improve the user experience, not only ensure the user security, but also improve software friendly degree.

Keywords CAPTCHA · Ajax · User experience

1 Introduction

Currently, CAPTCHAS are essentially challenge-response tests that are used to distinguish whether a user is a human or a computer. To date, numerous CAPT-CHA schemes have been proposed and deployed on various websites to secure online services from abuse by automated programs [1]. The hackers use the software to automatically register, send large quantities of spam to crack user passwords with brute-force method to steal user confidential and other increasingly serious problem. In order to ensure the stable operation of the user information

X. Chen (✉) · X. Lin
School of Science Anhui University of Science and Technology, 232001 Huainan, People's Republic of China

security and server, website CAPTCHA technology commonly used [2]. For deal with the CAPTCHA cracking techniques, CAPTCHA technology becomes more complex, it also gives the inconvenience of user visually recognize. User's network authentication code itself feel a burden, after visually recognizable, if enter CAPTCHA is wrong to give users great displeasure emotional [3]. This paper take advantage of the asynchronous transfer characteristics of Ajax technology and convenience do not need to refresh the page after the user enter the CAPTCHA immediately given judgment prompts the user input is correct or not. For avoiding the user clicks on the "Login" or "OK" button to submit information only after the return from the server to verify the success of the information. Using Ajax technology design graphics CAPTCHA can shorten the time that users wait for verification, avoid phenomenon CAPTCHA user misjudgment graphics, improved work efficiency, the use of Ajax technology design graphical validation code value.

2 CAPTCHA and Ajax

2.1 Captcha

A CAPTCHA is a program that protects websites against bots by generating and grading tests that humans can pass but current computer programs cannot [4]. For example, humans can read distorted text as the one shown below, but current computer programs can not.

2.2 Formulas

Ajax is a client script and it can interact the Web server the client-side Web development technology [5]. Fundamentally, the idea of Ajax is to improve the responsiveness and usability of Web pages by interacting with the server in a more flexible manner than the "page" model of standard HTML and HTTP. In most Web browsers, asynchronous communications with the Web server can be realized by the XMLHttpRequest object of the JavaScript. When each communication process completes, part of the Web page is updated dynamically based on the received data [6].

3 Enhance Website Interactivity Through Ajax Technology

At present, CAPTCHA cracking technology is also in constant development, in the past time, the researchers put forward a lot of identification method [7]. Generally, can be divided into three categories: based on template matching method, based on the character structure method and based on the neural network method [8]. In order to deal with the computer auto cracked, verification code also increasingly complex, in the artificial recognition, users also have difficult. Some website verification code examples as shown in Fig. 1 shows:

Complex CAPTCHAS bring inconvenience user login registration, verification code itself is of no value to the input information, just as a safety consideration, users in the input information, determine the input validation code, if the mis-carriage of justice lead to input error, in click confirm or login button, the web page prompt auth code error, the user is a major blow [9].

Users input user name information in the Web page form, the mouse left, triggering form on change events, data information to Ajax engine to handle, with JavaScript programming to realize Ajax engine function. Ajax engine instantiates a XMLHttpRequest object, Through this object to a server user input validation code strings, and back to the input string with reserve information than the results. So in the user complete input can know after input validation code is correct or not as shown in Fig. 2:

Fig. 1 Widely used CAPTCHA of some websites

Fig. 2 Structure of login system based on Ajax and CAPTCHA
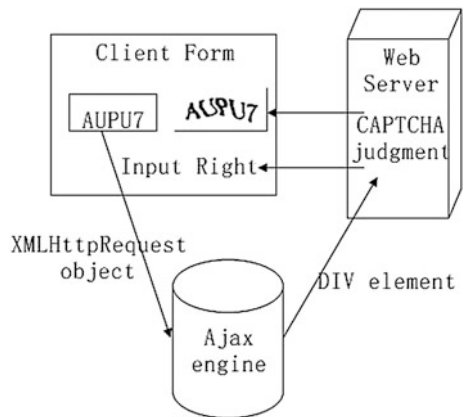
4 CAPTCHA Based on Ajax Prompt Realization

4.1 Design and Implementation of Input Form

Code of form which the user input information as below:

```
<form action = "dologin.jsp" method = get>
<input name = "captchas" onchange = "getPrompt()">
<div id = "txtPrompt" > </div>
</form>
```

In the form of user input information, add on change event to input CAPTCHA text box, when the user input string change, getPrompt function will be called, the function function is asynchronous transfer the user input to the server to judge the input string is correct or not. This layer div:txtPrompt, accept the return verification information from the server.

4.2 Use JavaScript Programming for Creating an Object Function

Code of getPrompt function as below:

```
<script type = "text/javascript">
    function getPrompt(){
    var xmlhttp;
    if (str =="")
  {document.getElementById("captchas").innerHTML = " ";
 return;}
    if(window.XMLHttpRequest){//for IE7+
      xmlhttp = new XMLHttpRequest();}
      else{//for ie5,ie6
      xmlhttp = new
ActiveXObject("Microsoft.XMLHTTP");}
      xmlhttp.onreadystatechange = function(){
      if(xmlhttp.readyState ==4 &&
xmlhttp.status ==200){
  document.getElementById("txtPrompt").innerHTML =
xmlhttp.responseText;
      }
      } xmlhttp.open("GET","checkCaptcha.jsp?q = " + str",true);
      xmlhttp.send();
  }
```

This function created XMLHttpRequest objects, transfer the user input validation code that assignment to parameter q and sent to the page :checkCaptcha.jsp.

Code of the page:checkCaptcha.jsp as below:

```
try {
        DBLink mydbLink = new DBLink();
        String  sql = "select  *  from  captcha  where  captchacode = '"
+ capcodes + "'";
          mydbLink.executeQuery(sql);
        ResultSet rs = mydbLink.getResultSet();
        while (rs.next()) {
        out.print("input right");
        }
        mydbLink.close();
    } catch (Exception e) {
        out.print(e);
    }
```

Judgment input string, if correct, send the string "input right" to this layer div:txtPrompt.

5 Conclusions

This paper analysis the characteristics of CAPTCHA and Ajax technology, find a way that improve user experience when user input CAPTCHA by using Ajax asynchronous transmission. Through the JavaScript language and JSP technology, a login system with CAPTCHA implemented. This System not only can ensure the security of the website, but also can improve software interactive friendly degree.

References

1. Chen X (2007) Miners positioning management system design and optimization. Energy Technol Management 6:33–34
2. Sheffer Y, Zorn G, Tschofenig H, Fluhrer S (2011) An EAP authentication method based on the encrypted key exchange (EKE) protocol. IETF RFC 6124
3. Babu R, Kumar KP, Rao S (2011) Implementation of secure multilayered CAPTCHA. IJAEST (6)

4. Fujino T (2007) SVG+Ajax+R: a new framework for WebGIS, computational statistics 22(4):511–520
5. Aboufadel E, Olsen J, Windle J (2005) Breaking the holiday inn priority club CAPTCHA. Coll Math J 36(2):101–108
6. Chellapilla K, Simard PY (2005) Using machine learning to break visual human. interaction proofs (HIPs)[C], In: Saul LK,Weiss Y, Bottou L, (eds) Advances in neural information processing systerms vol 17. MIT Press, Cambridge, pp 265–272
7. Huang S-Y, Lee Y-K, Graeme Bell Zhan-he Qu, 2008, A projection-based segmentation algorithm for breaking MSN and YAHOO CAPTCHAs [C]. In: Proceedings of the world congress on engineering, London
8. Hoeevar S PWNtcha[CP/OL], http://caca.zoy.org/wiki/PWNtcha
9. Son R, Tiwari BD (2010) Improved CAPTCHA method. Int J Comput Appli 1(25)

Numerical Characteristics of Rotor Angle in Power System with Random Excitations

Jianyong Zhang, Mingang Hua and Hongwei Wang

Abstract With the integration of renewable power and new load, much more random components are operating in the power system. In this paper, according to stochastic differential equation (SDE) theories, the SDE model of one machine and infinite bus (OMIB) system is constructed. The simulations show that nonlinear system can be replaced by linear system in the neighborhood of the initial point by means of Euler-Maruyama (EM) method. On the basis of the explicit solution of the linear stochastic differential equation, we have obtained the mathematical expectation, variance, and density function of rotor angle under nonlinear state. Moreover, the limit of the density function has been discussed as well.

Keywords SDE · Power system · Numerical characteristics · Density function

1 Introduction

The power network is a large and complicated system, which is often interfered by random components such as power system fault. With the integration of renewable power, such as wind energy, solar energy, ocean energy, and new loads (e.g., electrical vehicles) connected to the grid, the power system are brought into much

J. Zhang (✉)
Department of Mathematics and Physics, Hohai University, Changzhou 213022, China
e-mail: hohaizhangjy@hhu.edu.cn

M. Hua · H. Wang
College of Computer and Information, Hohai University, Changzhou 213022, China
e-mail: mghua@yahoo.cn

H. Wang
e-mail: visonhhuc@yahoo.cn

more random components. As a result, the probability characteristics of the power system with random components become new subjects.

The research on the effects of random components on the power system has won its initial success. The Ref. [1] gives the summery of the random components in power system and elaborates the research emphases of power system with different random components; Refs. [2, 3] have calculated the probabilistic stability of the system under the hypothesis that the initial values of the system are subjected to a certain distribution. Chen [4] has studied the first passage time problem of the dynamic power system under random disturbance with Hamilton theory; Zhang [5] treats the power absorption and injection after renewable power and electric vehicles connected to the grid as external random excitation to the power system, simulates the random excitation by the use of Gaussian white noise, and discusses the stability of a one machine and infinite bus system under random excitations. Qiu [6] discusses the voltage stability of power system under load fluctuations with stochastic differential equations. In Ref. [7], the authors studied linear power systems under sustained random perturbations, and obtained the necessary and sufficient conditions for almost sure asymptotic stability. As an example, a simulation was implemented in a one machine-infinite bus electric power system.

With the establishment of the stochastic differential equation model of a one machine infinite bus system, this paper expounds that the nonlinear model of the OMIB system can be replaced by its linear model in the neighborhood of the initial point without large errors by simulation, and offers the expression of explicit solutions to the OMIB linear differential equations. Moreover, on the theory of the explicit solutions to stochastic integral, the mathematical expectation, variance and density function in OMIB system are raised, by which the theoretical foundation for a further research is laid.

2 Stochastic Differential Equation Theories

2.1 Wiener Process and GAUSSIAN White Noise

Some stochastic differential equations can be constructed by order differential equations excited by Gaussian white noise, which are widely used in the field of natural science and engineering technology [8].

Suppose that $W(t)$ denotes the Gaussian white noise and $B(t)$ according to Ref. [9], the relationship between $W(t)$ and $B(t)$ can be expressed as

$$dB(t) = W(t)dt \tag{1}$$

2.2 *Itô Integral and its Properties*

Definition 1 *Itô* integral [10] refers to a special integral whose integral element is increment $dB(t)$ of Wiener process. It is generally denoted as $\int_0^a f(X(t), t) dB(t)$.

There are two import properties about *Itô* integral:
    Property 1 the mean value equals zero,

$$E[\int_0^a f(X(t), t) dB(t)] = 0 \tag{2}$$

Property 2 if $f(X(t), t)$ and $g(X(t), t)$ satisfy the conditions of *Itô* integral, then

$$E[\int_0^a f(X(t), t) dB(t) \int_0^a g(X(t), t) dB(t)] = \int_0^a f(X(t), t) g(X(t), t) dt \tag{3}$$

Specially, when $f(X(t), t) = g(X(t), t)$

$$E[\int_0^a f(X(t), t) dB(t)]^2 = \int_0^a f^2(X(t), t) dt \tag{4}$$

The two properties above are proved in detail respectively in Ref. [10].

2.3 Vector Stochastic Differential Equation and its Solution

Definition 2 Vector stochastic differential equation with Gaussian white noise can be written as

$$d\mathbf{X}(t) = \mathbf{f}(\mathbf{X}(t), t) dt + \mathbf{G}(\mathbf{X}(t), t) d\mathbf{B}(t) \tag{5}$$

Where $\mathbf{X}(t) = (X_1(t), X_2(t) \cdots, X_n(t))^T$ is the n-dimensional vector random variable; $\mathbf{B}(t) = (B_1(t), B_2(t) \cdots, B_n(t))^T$ is the n-dimensional Wiener process. The initial values $\mathbf{X}(t_0)$ and $\mathbf{B}(t)$ are independent of each other. If $\mathbf{f}(\mathbf{X}(t))$ and $\mathbf{f}(\mathbf{X}(t))$ are the linear functions of $\mathbf{X}(t)$, the linear stochastic differential equation can be obtained as follows

$$d\mathbf{X}(t) = (\mathbf{A}(t)\mathbf{X}(t) + \mathbf{a}(t)) dt + \mathbf{Q}(t) d\mathbf{B}(t) \tag{6}$$

Theorem 1 [9] *If Eq.* (5) *satisfies the following four conditions*:

1. $\mathbf{f}(\mathbf{X}(t), t)$ and $\mathbf{G}(\mathbf{X}(t), t)$, $(\mathbf{X}(t), t) \in [t_0, T] \times R^n$, are $\mathcal{L}^{n+1}$ measurable;
2. There exists a constant $K > 0$, such that for all $t \in [t_0, T], \mathbf{X}(t), \mathbf{Y}(t) \in R^n$

$$\| \mathbf{f}(\mathbf{X}(t),t) - \mathbf{f}(\mathbf{Y}(t),t) \|_2 \leq K \| \mathbf{X(t)} - \mathbf{Y}(t) \|_2;$$
$$\| \mathbf{G}(\mathbf{X}(t),t) - \mathbf{G}(\mathbf{Y}(t),t) \|_2 \leq K \| \mathbf{X(t)} - \mathbf{Y}(t) \|_2$$

3. There exists a constant $K > 0$, such that for all $t \in [t_0, T], \mathbf{X}(t), \mathbf{Y}(t) \in R^n$,

$$\| \mathbf{f}(\mathbf{X}(t),t) \|_2^2 \leq K^2(1+ \| \mathbf{X}(t) \|_2^2), \| \mathbf{G}(\mathbf{X}(t),t) \|_2^2 \leq K^2(1+ \| \mathbf{X}(t) \|_2^2)$$

4. $\mathbf{X}(t_0)$ is σ-algebra measurable, and $E(\| \mathbf{X}(t_0) \|_2^2) < \infty$

Then there exists only one strong solution $\mathbf{X}(t)$, $t \in [t_0, T]$ to stochastic differential Eqs. (5) or (6), and satisfies $\sup E(\| \mathbf{X}(t) \|_2^2) < \infty, t \in T.$ $\| \cdot \|_2$ Denotes norm, which refers to modular arithmetic and 2-norm for vector and matrix, respectively.

Theorem 2 [9] *When $A(t) = A$ in the linear stochastic differential Eq. (5), its solution process can be expressed as*

$$\mathbf{X}(t) = e^{\mathbf{A}(t-t_0)}\mathbf{X}(t_0) + \int_{t_0}^t e^{\mathbf{A}(t-s)}[\mathbf{a}(s)ds + \mathbf{Q}(s)d\mathbf{B}(s)] \tag{7}$$

3 Stochastic Differential Equations of OMIB

A one machine and infinite bus (OMIB) system is applied in the simulations. As shown in Fig. 1, the system consists of a generator, two transformers and an infinite bus. In the figure, the transformer reactance $x_{T1} = 0.138$ pu and $x_{T2} = 0.122$ pu; the line reactance $x_l = 0.243$ pu; the generator transient reactance $x_d' = 0.295$ pu; the inertia constant $M = 2{,}569.8288$ pu; and the damping coefficient $D = 2.0$ pu. The initial system power $P_0 = 1.0$ pu, $Q_0 = 0.2$ pu, voltages behind the reactance $E' = 1.41$ pu and rotor angle $\delta_0 = 34.46°$. Per unit system: $S_B = 220$ MVA, $U_{B(220)} = 209$ kV.

Under the deterministic circumstance [10, 11], the motion equation of OMIB system can be written as follows:

$$M\frac{d^2\delta}{dt^2} + D\frac{d\delta}{dt} = P_m - \frac{E'U}{X_\Sigma}sin\delta \tag{8}$$

Fig. 1 One machine and infinity bus systems
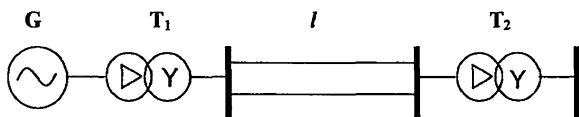
where δ is the rotor angle; ω is the rotating speed; P_m is the mechanical power; and E' is the internal voltage; $X_{\sum}$ is total reactance; U is the infinite bus voltage.

After adding a random excitation term to the right-hand side of the motion Eq. (8), the motion equation can be written as

$$\begin{cases} \frac{d\delta}{dt} = \omega - 1 \\ M\frac{d\omega}{dt} = P_m - \frac{E'U}{X_\Sigma}\sin\delta - D(\omega - 1) + \sigma_0 W(t) \end{cases} \tag{9}$$

Where, the initial values $\delta_0 = 34.46°, \omega_0 = 1$, the random excitation term $W(t)$ can be interpreted as random power fluctuations which can be induced by renewable energy generation, or caused by random load such as electric vehicle. σ_0 is the intensity of random excitation, and its magnitude presents a probable maximum value of the random excitation.

Suppose $\Delta\delta = \delta - \delta_0, \Delta\omega = \omega - \omega_0$, the linearized state Eq. of (9) can be written as

$$\begin{cases} \frac{d\Delta\delta}{dt} = \Delta\omega \\ \frac{d\Delta\omega}{dt} = -\frac{E'U\cos\delta_0}{MX_\Sigma}\Delta\delta - \frac{D}{M}\Delta\omega + \frac{\sigma_0}{M}W(t) \end{cases} \tag{10}$$

Its vector equation is as follows

$$d\mathbf{X}(t) = \mathbf{A}\mathbf{X}(t)dt + \mathbf{Q}dB(t) \tag{11}$$

where $\quad \mathbf{X}(t) = \begin{pmatrix} \Delta\delta \\ \Delta\omega \end{pmatrix}, \mathbf{X}(t_0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathbf{A} = \begin{pmatrix} 0 & 1 \\ a & b \end{pmatrix},$

$a = -\dfrac{E'U\cos\delta_0}{MX_\Sigma}, \ b = -\dfrac{D}{M}, \ \mathbf{Q} = (0, \sigma_0/M)^T.$

According to (7), we can obtain the explicit solution to the Eq. (11)

$$\mathbf{X}(t) = e^{\mathbf{A}(t-t_0)}\mathbf{X}(t_0) + \int_{t_0}^{t} e^{\mathbf{A}(t-s)}\mathbf{Q}dB(s) \tag{12}$$

If the random excitation intensity σ_0 equals 0.04, comparison of two rotor angle curves can be made by applying EM method [12] to linear and nonlinear stochastic differential equations, respectively. The comparison results are illustrated in Fig. 2.

With comparing the two curves, we find that the rotor angle of the linear model is almost identical to the nonlinear model under small random excitation. So the linear model can be used to analyze the numerical characteristics of nonlinear power system under small random excitations.

In the Ref. [11], we can see the same method was used to study the local stability of power system, and the author has provide a detailed description about this method as well.

Fig. 2 Comparison of linear and nonlinear models
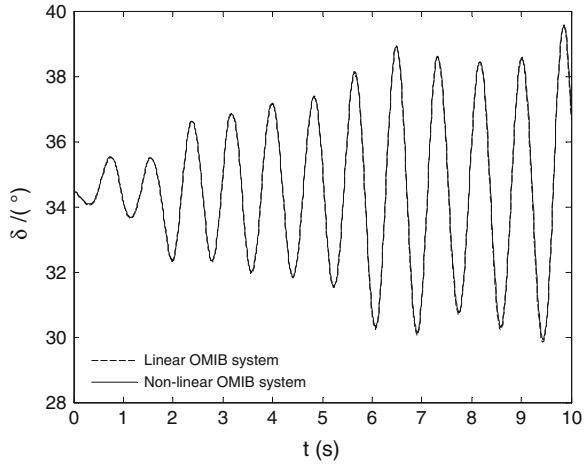
4 Numerical Characteristics

Firstly, we calculated the matrix exponential function of A

$$[sI - A]^{-1} = \begin{bmatrix} s & -1 \\ -a & s-a \end{bmatrix}^{-1} = \frac{1}{s^2 - bs - a}\begin{bmatrix} s-b & 1 \\ a & s \end{bmatrix}$$

Suppose $\alpha = -b/2, \beta = \sqrt{-a - b^2/4}$, then

$$e^{At} = L^{-1}[(sI-A)^{-1}] = e^{-\alpha t}\begin{bmatrix} \cos \beta t - \frac{\alpha}{\beta}\sin \beta t & \frac{1}{\beta}\sin \beta t \\ \frac{\alpha}{\beta}\sin \beta t & \cos \beta t + \frac{\alpha}{\beta}\sin \beta t \end{bmatrix}$$

Since $X_0 = [0,0]^T$, according (12),

$$X(t) = [X_1(t), X_2(t)]^T = [\Delta\delta, \Delta\omega]^T = \int_0^t e^{A(t-s)}QdB(s)$$

$$= \int_0^t \frac{1}{\beta}e^{-\alpha(t-s)}\begin{bmatrix} \beta\cos\beta(t-s) - \alpha\sin\beta(t-s) & \sin\beta(t-s) \\ \alpha\sin\beta(t-s) & \beta\cos\beta(t-s) + \alpha\sin\beta(t-s) \end{bmatrix}\begin{bmatrix} 0, \frac{\sigma^0}{M} \end{bmatrix}^T dB(s)$$

Because the rotor angle δ plays an important role in the power system, we only analyze $X_1(t)$

$$X_1(t) = \Delta\delta = \frac{\sigma_0}{M}\int_0^t \frac{1}{\beta}e^{-\alpha(t-s)}\sin\beta(t-s)dB(s)$$

The numerical characteristics of $X_1(t)$ are as follows,

According to Eq. (2), we calculate the mathematical expectation

$$m(t) = E(X_1(t)) = 0$$

According to Eq. (4), we calculates the variance

$$\sigma^2(t) = E\left[(X_1(t) - EX_1(t))(X_1(t) - EX_1(t))^T\right]$$

$$= E\left[\left(\frac{\sigma_0}{M}\int_0^t \frac{1}{\beta} e^{-\alpha(t-s)} \sin\beta(t-s) dB(s)\right)\left(\frac{\sigma_0}{M}\int_0^t \frac{1}{\beta} e^{-\alpha(t-s)} \sin\beta(t-s) dB(s)\right)^T\right]$$

$$= -\frac{\sigma_0^2 \sin^2\beta t}{2\alpha\beta^2 M^2 e^{2\alpha}} + \frac{\sigma_0^2}{2\alpha\beta(\alpha^2 + \beta^2)M^2}\left[\frac{\beta}{2}(1 - e^{-2\alpha t}\cos 2\beta t) - \frac{\alpha\sin 2\beta t}{2} e^{-2\alpha t}\right]$$

According to Ref. [9], we can know that $X_1(t)$ is still Gaussian random process, and density function is determined by its mathematical expectation and variance. The density function of $X_1(t)$

$$f(X_1, t|X_{10}, 0) = \frac{1}{\sqrt{2\pi}|\sigma^2(t)|^{\frac{1}{2}}} \exp\left(-\frac{[X_1 - m(t)]^2}{2\sigma^2(t)}\right)$$

The mathematical expectation, variance and density function of $\Delta\delta(t)$ we have discussed above are under linear state. As for nonlinear state variable $\delta(t)$,

$$E[\delta(t)] = E[\Delta\delta(t) + \delta_0] = m(t) + \delta_0 = \delta_0$$
$$D[\delta(t)] = D[\Delta\delta(t) + \delta_0] = \sigma^2(t) + D(\delta_0) = \sigma^2(t)$$
$$f(\delta, t|\delta_0, 0) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2(t)}} \exp\left(-\frac{[\delta - \delta_0]^2}{2\sigma^2(t)}\right)$$

When $t \to \infty$, since $\alpha = -b/2, \beta = \sqrt{-a - b^2/4}$,

$$\lim_{t\to\infty} E[\delta(t)] = \delta_0 \tag{13}$$

$$\lim_{t\to\infty} D[\delta(t)] = \lim_{t\to\infty} \sigma^2(t) = \frac{\sigma_0^2}{2abM^2} \tag{14}$$

$$\lim_{t\to\infty} f(\delta, t|\delta_0, 0) = \sqrt{\frac{abM^2}{\pi\sigma_0^2}} \exp\left(-\frac{abM^2}{\sigma_0^2}(\delta - \delta_0)^2\right) \tag{15}$$

According to Eqs. (13), (14), the limits of mean value $m(t)$ and variance $\sigma^2(t)$ of rotor angle δ are both bounded, i.e. the power angle δ is mean stable and mean square stable, which is not only accordant to the conclusion of the Ref. [5] but also further complement it in the perspective of nonlinearity. The limit distribution in (15) is a normal distribution with mean δ_0 and variance $\delta_0^2/(2abM^2)$.

5 Conclusion

In this paper, taking the power fluctuations caused by the integration of renewable power and new load to the grid as the small Gaussian Random excitation, we have constructed the stochastic differential equations of a one machine and infinite bus system and provided the explicit solutions to linear system in the neighborhood of initial points. Furthermore, because of the important role power angle plays in power system, we have analyzed its mathematical expectation and variance under random small excitation, from which its density function has been obtained.

Simultaneously, the limits of mathematical expectation and variance of rotor angle indicate that the system is mean stable and mean square stable under small random excitations, which is not only accordant to the conclusion of the Ref. [5] but also further complement it in the perspective of nonlinearity. The limit condition of density function indicates that the system will show more obvious characteristics of normal distribution under long-term small random excitations, by which the theoretical foundation for the subsequent discussion on its characteristic functions and power spectrum analysis, etc., is laid.

References

1. Ju P, Qin C, Huang Y et al (2012) Research trends of power system modeling geared to smart grid. Autom Electr Power Syst 36:1–6
2. Ju P, Wu GY, Li Y (1991) Fundamental theorems on probabilistic stability of power system. Proc CSEE 11:17–25
3. Wang KW, Chung CY, Tse CT et al (2000) Improved probabilistic method for power system dynamic stability studies. IEE Proc-Gener Transm Distrib 147:37–43
4. Chen LC, Zhu WQ (2010) First passage failure of dynamical power systems under random perturbations. Sci China Technol Sci 53:2495–2500
5. Zhang JY, Ju P, Yu YP et al (2012) Responses and stability of power system under small gauss type random excitation. Sci China Technol Sci 55:1873–1880
6. Qiu Y, Zhao J, Zhu Y (2009) Effect of stochastic load disturbance on power system voltage stability based on bifurcation theory. Electr Power Autom Equip 29:77–81
7. Humberto V, Luis V, Wolfgang K (2012) Stability of linear stochastic systems via Lyapunov exponents and applications to power systems. Appl Math Comput 218:11021–11032
8. Zhang BG, Zhao YZ (1998) Stochastic differential equations in science and engineering. Ocean Press, Beijing
9. Henderson D, Plaschko P (2006) Stochastic differential equations in science and engineering. World Scientific Publishing, Singapore
10. Kundur P (1994) Power system and control. McGraw-Hill, New York
11. Higham DJ (2001) An algorithmic introduction to numerical simulation of stochastic differential equations. SIAM Rev 43(3):525–546

An Epidemic-Dynamics-Based Model for CXPST Spreading in Inter-Domain Routing System

Yu Wang, Zhenxing Wang and Liancheng Zhang

Abstract We study the CXPST attack which aims at the destruction of inter-domain routing system and propose a spreading model to represent the threatening scale. By analyzing the process we illuminate the mechanism of how CXPST seizes the BGP deficiencies to paralyze the Internet control plane from the traffic attack of data plane, and then the spreading model named as EDM-CS is presented based by the epidemic dynamics theory. Parameters of the model are closely associated with the real network topology and BGP router overloading condition which reflect the features of the CXPST spreading. In virtue of the classical BA scale-free network, spreading density that derives from EDM-CS behaves great consistency with the simulation results based on the collected data from CAIDA. This model can help understanding CXPST attack and providing a basis for predicting the spreading trend, as well as investigating effective defense strategy.

Keywords CXPST · Inter-domain routing system · Spreading model · Epidemic dynamics

Y. Wang (✉) · Z. Wang (✉) · L. Zhang
National Digital Switching System Engineering & Technological Research Center, 450002 Zhengzhou, China
e-mail: stonchor@gmail.com

Z. Wang
e-mail: Zhenxingwang@gmail.com

L. Zhang
e-mail: A1ien@gmail.com

1 Introduction

As the most significant infrastructure of nowadays Internet, inter-domain routing system is the key connection between ASes (anonymous systems) which not only have to transmit the data stream but also need to exchange the network reachability information. For ISPs utilize this system to make strategies of access control generally, the topological structure and security features of inter-domain routing system deeply influence the functions of the whole Internet.

The scale of inter-domain routing systems is expanding with the new ISPs emerging rapidly, and the establishments of new multiform connections further accelerate the complexity of the inter-domain routing system. Frequent security events [1, 2] towards inter-domain routing system have happened in recent years, more complicated measures which cause more loss, urging researchers to recognize the fundamentality and severity of these problems.

Schuchard proposed a sharp attack against inter-domain routing system, named as (CXPST) Coordinated cross plane session termination [3], which attempts to destruct the whole Internet control plane through the low-rate DDoS onto the data plane. The simulations demonstrate that this method can potentially spread and disrupt the inter-domain routing system and almost the whole Internet, which even needs several days to fully recover.

However, few works have investigated how the CXPST spread and pervade into the inter-domain routing system by now, which results in many obstacles to understanding the CXPST-like attacks and improving the protection of inter-domain routing system. In this paper, an epidemic-dynamics-based model for CXPST spreading (EDM-CS) is presented by the inspiration of epidemic theory. Our objective is to provide a methodology to depict the spreading mechanism of CXPST attack, based on which, operators can study the cascading failure phenomenon of inter-domain routing system caused by CXPST-like attacks, as well as the prevention or recovery methods against the attack can be further investigated.

2 EDM-CS Model

The process of CXPST is similar with epidemic spreading. (1) In terms of the structure of target, their objects are inter-domain routing system and the human society respectively, and both of them are complex giant systems composed of many interconnected and interacted sub-systems. (2) In terms of the spreading course, their initial attack areas are both a tiny parts of the system, and it is by controlling the attacking intensity and frequency that to spread threat gradually into the global system. (3) In terms of the susceptive unit, the human body owns the ability of self-healing, and the BGP router also has the ability to recover from processing capability decline of attack. (4) In terms of the pathogen, BGP

UPDATE message and the virus are chosen respectively to cause the target capacity down till the failure.

Epidemic dynamics theories have been used by researchers to study and model the spreading of threats on the Internet [4–6], like virus, worm and even outbreak disasters, helping people to understand the spreading phenomena [7]. Thereby to investigate the effect of CXPST to the Internet inter-domain routing system, we establish a CXPST spreading model in virtue of the epidemic dynamic theory. First, the related components will be defined in Sect. 2.1, and then followed by the description of the EDM-CS Model.

2.1 Definitions

Definition 1 Present an unweighted and undirected graph $G = <V, E>$ as the inter-domain routing system. $\forall v \in V$, v denotes an AS; $\forall e \in E$, e is an unordered pair of elements in V to represent an interconnection between ASes.

Definition 2 Every BGP router in the inter-domain routing system is susceptive to CXPST attack. Let U denotes the set of routers in the inter-domain routing system. A denotes the set of nodes under attack and N is the set of normal functioning nodes, then $A \cup N = U$ and $A \cap N = \phi$

Definition 3 The density of infected nodes $\rho_n(t)$ is used to describe the ratio of attacked nodes with n degree to the total nodes, denoting the probability that nodes with degree n under CXPST attack, where $0 \leq \rho_n(t) \leq 1$.

Definition 4 The BGP router linking to the attacked routers turns into the Attacked state with probability w, where $0 < w \leq 1$. While the node suffering from attack endeavors to process the stacked UPDATE messages and resume to Normal state with probability r, where $0 \leq r \leq 1$.

Definition 5 Let λ be the actual spreading intensity, $\lambda = w/r$. As we could set $r = 1$, thus $\lambda = w$, while this presupposition will only influence the definition of CXPST spreading timing scale.

Definition 6 For a path on which CXPST attack spreads, the router under attack at an early time is defined as pre-node, and the router influenced by the pre-node afterwards is called the next-node.

2.2 Model for CXPST Spreading

According to the degree of the router within inter-domain routing system, a single Mean-Field reaction equation for the density of attacked nodes while CXPST spreading is proposed:

$$\frac{\partial \rho_n(t)}{\partial t} = -\rho_n(t) + \lambda n[1 - \rho_n(t)]\left[\langle n \rangle^{-1} \sum_{i=1}^{k} iP(i)\, \rho_i(t)\right] \tag{1}$$

In Eq. (1) all the higher order corrections in $\rho_n(t)$ have been ignored for $\rho_n(t) < 1$. The first term $-\rho_n(t)$ indicates the routers suffered from attack resume normal state with unit rate by themselves. The second term denotes the average probability density of newly attacked router generated by every attacked router. This term is proportional to the (1) actual spreading rate λ (2) degree of the router n (3) the probability of being normal before time t, denoted as $[1 - \rho_n(t)]$ (4) the probability that any link of this node points to a normal router, namely $\left[\langle n \rangle^{-1} \sum_{i=1}^{k} iP(i)\rho_i(t)\right]$.

In respect that the probability that a router with n degree suffers from CXPST attack is $\rho_n(t)$, the probability that it being a normal node is $[1 - \rho_n(t)]$.

Let $K(\lambda, P) = \langle n \rangle^{-1} \sum_{i=1}^{k} iP(i)\rho_i(t)$, where k denotes the max degree of nodes in the system, $\langle n \rangle$ is the average degree, $P(i)$ indicates the ratio of nodes with degree i to the whole nodes, thus $K(\lambda, P)$ denotes the probability that a router links to an attacked router.

Impose $\left\{-\rho_n(t) + \lambda n[1 - \rho_n(t)]\left[\langle n \rangle^{-1} \sum_{i=1}^{k} iP(i)\rho_i(t)\right]\right\} = 0$, we obtain

$$\rho_n(\lambda, P) = \frac{\lambda n K(\lambda, P)}{1 + \lambda n K(\lambda, P)} \tag{2}$$

From Eq. (2) we observe that along with the increase of spreading intensity, the probability that routers with any degree that suffer from attack will get higher, and $K(\lambda, P)$ indicating the probability that points to a attacked node through any link will get higher as well.

3 Assessment

Similar to the epidemic spreading, the EDM-CS model for CXPST is highly dependent on the network topology. From Eq. (2) we could not directly obtain the CXPST spreading situation within inter-domain routing system, and there in no standardized model for the inter-domain routing system. Here, existing network topology models are utilized to stimulate the inter-domain routing system, as to analyze the spreading features of EDM-CS model.

It is observed that for the nodes with i degrees, the probability ρ_i indicating the node being attacked is mainly influenced by the degree-correlated attributes including the average degree, the maximal degree and the ratio of nodes with degree i to the whole nodes.

There are remarkable characteristics of Power-Law for the inter-domain routing system. Faloutsos [8, 9] studied and illuminated the topological characteristics of the AS-level Internet who pointed out multi Power-Laws relationships such as the

degree of node, rank exponent and hop-plot exponent. Therefore, we would like to adopt classical Barabasi-Albert (BA) scale-free network [10] whose growth and preferential attachment characteristics can well depict the Power-Laws of inter-domain routing system, to help analyzing and assessing the EDM-CS model.

Define m as the minimal number of degree for each node in BA network, thus the average connectivity $\langle n \rangle = 2m$, and the full connectivity distribution $P(n) = 2m^2 n^{-3}$, yielding the relation

$$\mathrm{K}(\lambda, P) = m\lambda \mathrm{K}\,(\lambda,\ P) \int_m^\infty \frac{dn}{n(1 + n\lambda \mathrm{K}\,(\lambda, P))} \tag{3}$$

From which we obtain

$$\mathrm{K}\,(\lambda,\ P) = \frac{e^{-1/m\lambda}\left(1 - e^{-1/m\lambda}\right)^{-1}}{\lambda m} \tag{4}$$

As the time tends to be infinite, let ρ denote the steady density of attacked nodes, combine with Eq. (2) we get

$$
\begin{aligned}
\rho &= \sum_n P(n)\rho_n = \sum_n \left(2m^2 n^{-3}\right) \frac{\lambda n \mathrm{K}(\lambda,\ P)}{1 + \lambda n \mathrm{K}(\lambda,\ P)} \\
&= 2m^2 \lambda \mathrm{K}(\lambda,\ P) \int_m^\infty \frac{dn}{n^2[1 + n\mathrm{K}(\lambda, P)]} \\
&= 2m^2 \lambda \mathrm{K}(\lambda,\ P)\left[\frac{1}{m} + \lambda \mathrm{K}(\lambda,\ P)\ln\left(1 + \frac{1}{m\lambda \mathrm{K}(\lambda,\ P)}\right)\right]
\end{aligned}
\tag{5}
$$

From here the expression for ρ is finally

$$\rho \propto e^{-1/m\lambda} \tag{6}$$

It is illuminated from Eq. (6) that under given m in the network, when the spreading density λ is increasing, the ratios of attacked nodes with different degrees suffering from CXPST attack will be higher in the steady state.

4 Simulation

4.1 Data and Configuration

There are several publicly available sources of raw and processed BGP data. To validate the proposed spreading model, parts of the AS-level topology provided by CAIDA [11] project are adopted as the data source. Using the same BGP ISS [12] in our previous work, the main information obtained is shown in Table 1, in which 30,192 ASes and 811,076 links are selected to form the undirected graph for simulation.

Table 1 Information of selected AS-level topology

# of ASes	# of Links	Max degree	Min degree	Average degree
30192	811076	2526	1	4.72

Deploying a large-scale botnet is a necessary precondition to launch CXPST, adversaries in control of the botnet selectively disrupts BGP sessions to generate a surge of BGP updates which overwhelm the capacity of routers and disturb the routing decisions.

From the analysis in Sect. 3, there are three factors to be considered for the deployment of botnet. (1) Distribution of bots, the bot placement will affect DDoS traffic control to the key link of the topology, and we utilize the Waledac [13] data set to construct botnets model of bots distribution. (2) Total amount, the quantity of the bots will influence the attack intensity heavily, and the amount which ranges from 12 to 200 K of bots will be used in the simulation. (3) Bandwidth, CXPST force the key link cut and connected by low-rate DDoS, therefore the bandwidth is set at the range of 512 K–1 M bps.

4.2 Evaluations

4.2.1 Comparison

In order to validate the EDM-CS model, we first perform the analytical prediction based on the EDM-CS model under BA networks, then implement simulative CXPST attack to the real AS-level Internet topology and calculate the density of attacked nodes with different degrees of intensity, the two parts of result are compared and analyzed.

It is observed from Eq. (6) that ρ is not negative and only when $\lambda = 0$ that $\rho = 0$, indicating absence of any spreading threshold or critical point in this model. Once the spreading intensity $\lambda > 0$, the CXPST attack could spread and maintain in a balanced state, showing the vulnerability of BA networks to CXPST. In this case, attacks performed with high spreading intensity can result in extremely high prevalence level to the inter-domain routing system based on the attacks to a few links and nodes initially.

Our numerical study between the theoretical prediction from EDM-CS model and experimental result from practical simulation is reported in Fig. 1, where the average degree is set with 4.72 when calculating under BA networks, which is consistent with the simulation topology. In Fig. 2, we plot the relationships between the attack intensity λ and spreading density ρ, it is observed that the density of attacked nodes is increased with the growth of the spreading density. The max spreading density is approaching to 0.7 with the average degree of 4.72, which is enough to paralyze the whole inter-domain routing system. We observe

Fig. 1 Density of attacked nodes ρ as the function of λ in BA networks (*full line*) and the simulation topology from CAIDA (*dotted line*)
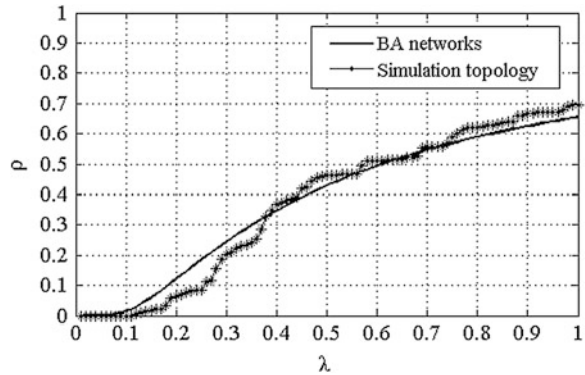
that the dotted line indicating the simulation result is close to the full line, which demonstrate that the EDM-CS model can reflect the spreading capability of CXPST attack.

4.2.2 Influence of Degree

After the validation for the EDM-CS model in respect of correctness, here we examine the influence of node degree on the spreading scale under the given BA scale-free topology.

The relationships between attack intensity λ and spreading density ρ in case of different average degrees are shown in Fig. 2. It is observed that the density of attacked nodes is increasing while the spreading intensity is upgraded; besides, under the identical attack intensity, the higher the average degree of the BA network is, the larger the proportion of attacked nodes is.

As degree distribution of inter-domain routing system is in conformity to Power-Laws, this extremely heteropical characteristic indicates that most of the nodes possess less degree while few key nodes have high degree. The higher the

Fig. 2 Influences of different average degrees on the density of attacked nodes
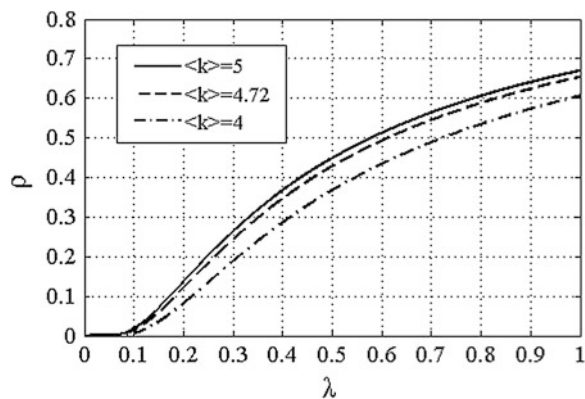
average degree of scale-free network is, the higher the degree of key nodes is, which means when these nodes or correlated links are threatened by the intentional CXPST-like assault, the whole networks are more fragile and the scale of attacked nodes is larger. Accordingly, Fig. 2 illustrates that the EDM-CS model can depict the influence of degree distribution on the CXPST spreading.

5 Conclusion

As one of the most destructive assaults to the inter-domain routing system in recent years, CXPST has drawn a lot of attentions. This paper proposes the EDM-CS model to depict the cascading spreading of CXPST attack based on the analysis of the operational process and functioning mechanism, and two metrics including density of attacked nodes and spreading intensity are introduced to evaluate the influence of botnet on the attack spreading. Combining with the consideration of the topological characteristics of the inter-domain routing system, AS-level topology from CAIDA are used to simulate the spreading of CXPST via the inter-domain routing system. Simulation results show that the theoretical prediction from EDM-CS is in accordance with the practical simulation to a great extent which indicates that the model has high veracity and practical worthiness. Moreover, we observe that the influence of average degree of network is also conformed to the situation of CXPST spreading scale and the Power-Law fragility of inter-domain routing system.

References

1. Renesys (2008) Pakistan hijacks YouTube. http://www.renesys.com/blog/2008/02/pakistan_hijacks_youtube_1.shtml
2. Wan T, Van Oorschot P C (2006) Analysis of BGP prefix origins during Google's May 2005 outage. Proceedings of the 20th international parallel and distributed processing symposium. 135–142
3. Schuchard M, Vasserman E Y, Mohaisen A, Kune D F, Hopper N, Kim Y (2010) Losing control of the internet: Using the data plane to attack the control plane. Proceedings of the 17th ACM conference on computer and communication security. 726–728
4. Li X, Wang X (2006) Controlling the spreading in small-world evolving networks stability, oscillation, and topology. IEEE Trans Autom Control 3(51):534–540
5. Zou CC, Towsley D, Gong W (2006) On the performance of Internet worm scanning strategies. Perform Eval 63:700–723
6. Chung-Yuan H, Yu-Shiuan T, Tzai-Hung W (2010) A Network-based simulation architecture for studying epidemic dynamics. Simulation 86(5–6):351–368
7. Ogierman A, Elsasser R (2012) The impact of the power law exponent on the behaviour of a dynamic epidemic type process. Proceedings of 24th ACM symposium on parallelism in algorithm and architectures, 131–139
8. Faloutsos M, Faloutsos P, Faloutsos C (1999) On power-law relationships of the Internet topology. ACM SIGCOMM Comput Commun Rev 4(29):251–262

9. Siganos G, Faloutsos M, Faloutsos P, Faloutsos C (2003) Power laws and the AS-level internet topology. IEEE/ACM Trans Networking 4(11):514–524
10. Barabasi A, Albert R (1999) Emergence of scaling in random networks. Science 289(5439):509–512
11. Routeviews (2012) BGP AS links. http://as-rank.caida.org
12. Guo Y, Wang Z, Luo S, Wang Y (2011) A cascading failure model for inter domain routing system. Int J Commun Syst 25:1068–1076
13. Sinclair G, Nunnery C, Kang B B (2009) The waledac protocol: The how and why. Proceedings of the 4th IEEE international conference on malicious and unwanted software, 69–77

An Approach for Diversity and Convergence Improvement of Multi-Objective Particle Swarm Optimization

Shan Cheng, Min-You Chen and Gang Hu

Abstract To improve the diversity and convergence of multi-objective optimization, a modified Multi-Objective Particle Swarm Optimization (MOPSO) algorithm using Step-by-step Rejection (SR) strategy is presented in this paper. Instead of using crowding distance based sorting technique, the SR strategy allows only the solution with the least crowding distance to be rejected at one iteration and repeat until the predefined number of solutions selected. With introduction of SR to the selection of particles for next iteration, the modified algorithm MOPSO-SR has shown remarkable performance against a set of well-known benchmark functions (ZDT series). Comparison with the representative multi-objective algorithms, it is indicated that, with SR technique, the proposed algorithm performs well on both convergence and diversity of Pareto solutions.

Keywords Multi-objective optimization · Particle swarm optimization · Crowding distance · Step-by-step rejection

1 Introduction

Multi-objective optimization (MO) has been widely studied in view of the practical importance in the field of science and engineering. For MO, the aim is to identify Pareto optimal set. And generally, for a competitive MO algorithm, two of

S. Cheng (✉) · M.-Y. Chen · G. Hu
State Key Laboratory of Power Transmission Equipment and System Security and New Technology, Chongqing University, Chongqing 400044, China
e-mail: hpucquyzu@cqu.edu.cn

M.-Y. Chen
e-mail: minyouchen@cqu.edu.cn

G. Hu
e-mail: gang.hu.cq@gmail.com

the main issues [1] to be considered are how to make the Pareto front produced by the algorithm converge to the true Pareto front (if known) and how to promote the diversity of Pareto solutions of smooth and uniform spreading. To address the second one, NSGA-II crowding distance (CD) based sorting [2] is widely used to filter certain solutions from the set these solutions belong to. The basic operations are computing the CD of each solution in the set, sorting the solutions according to CDs and selecting a predefined number of solutions with relatively larger CDs. The direct "selection" ignores the effect on the CD(s) of the selected particles caused by the nearby obsolete solution(s), which makes the selected solutions too sparse and decreases the diversity of the solutions.

To avoid negative effects on diversity caused by this direct selection, a step-by-step rejection (SR) is proposed, through which well-distributed solutions can be identified from a given set. With introduction of SR and on the basis of prophase studies [3, 4], an improved multi-objective particle swarm optimization (MOPSO-SR) is presented, whose performance has been verified on a set of benchmark functions and compared with commonly recognized effective MO algorithms.

2 A Step-by-Step Rejection Based Multi-objective PSO

As a bio-inspired technique, PSO has been extended for MO. Several variants of the PSO algorithm have been presented to deal with MO problems based on the concept of Pareto optimality, in which the NSGA-II CD based sorting [2] technique was popularly used to promote the diversity of Pareto solutions. In prophase study, Adaptive Weighted PSO (AWPSO) [3] and Adaptive Evolutionary PSO (AEPSO) [4] employed random selection method and NSGA-II CD based sorting technique to select the particles for next iteration from the current non-dominated solutions. To avoid the drawback of CD based sorting technique and to improve the diversity of Pareto solutions while keeping good convergence, on the basis of AWPSO and AEPSO, we present a SR based multi-objective PSO (MOPSO-SR).

2.1 Stem-by-Step Rejection Strategy

Combine the current population P (including N particles) and its offspring population Q to generate a new population R, a series of non-dominated solution sets with different levels are obtained according to non-domination. If a certain number of non-dominated solutions needs to be selected from the set theses solutions belong to, the NSGA-II CD based sorting technique is employed for diversity preservation. As Fig. 1 illustrated, five solutions needs to be chosen from S_1, S_2, …, S_{10}. Assume the CDs of S_1 and S_{10} to be ∞, the others' are listed in Table 1. The five solutions S_1, S_7, S_8, S_9 and S_{10}, whose CDs are larger than the others', are selected directly, while the rest are rejected. However, this approach makes S_7, S_8
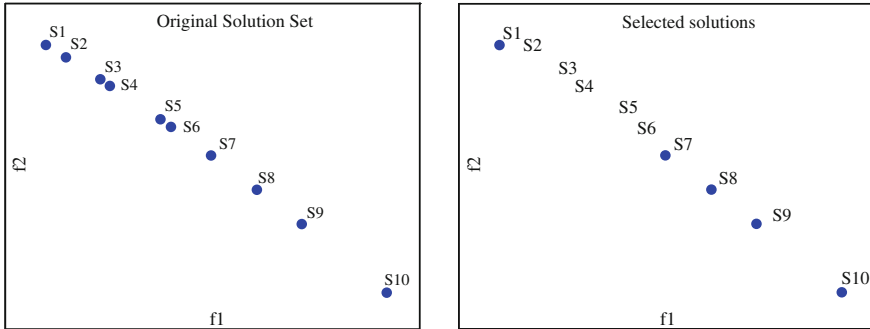
Fig. 1 Example of applying NSGA-II non-dominated sorting technique

Table 1 CDs of S_2, S_{83}, ... S_9

	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9
CD	0.299	0.243	0.337	0.346	0.294	0.501	0.539	0.800

and S_9 rather crowded and S_1 and S_7 rather sparse. It's clearly that the distribution of selected solutions using such selection method isn't good.

NSGA-II CD based sorting technique is a direct selection approach. Different from the scheme of direct selection, we advocate realizing the selection via SR. SR strategy allows only one solution with the least CD to be rejected at each step, and via SR, well distributed solutions of good diversity can be obtained. Firstly, compute the CDs of all non-dominated solutions, and reject the solution with the least CD; then re-compute the CDs of the rest non-dominated solutions and reject the solution with the least CD. Circulate until the predefined number of solutions is left. Figure 2 illustrates selecting five solutions from the ten solutions shown in Fig. 1 using the SR strategy, in which S_3, S_6, S_2, S_8 and S_5 are rejected in turn. Compared with CD based sorting technique, the selected solutions are of even distribution and good diversity. Therefore, SR strategy, which realizes the selection through SR, is more beneficial to the improvement of diversity.

2.2 SR Applied in the Selection of Particles for the Next Iteration

SR is applied to selection of particles for the next iteration. The number of non-dominated solutions (N_{ND}) identified from the population R can be larger or less than N. Regarding the selection of the particles for next iteration, when $N_{ND} > N$, generally, particles can be selected from the non-dominated solutions by random selection [3, 5] or NSGA-II CD based sorting [4, 6]. Because of the disadvantage
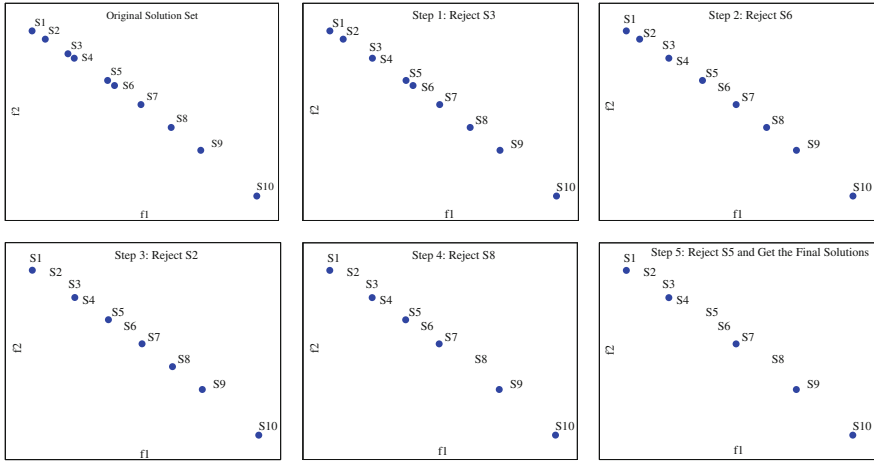
Fig. 2 Example of applying SR strategy
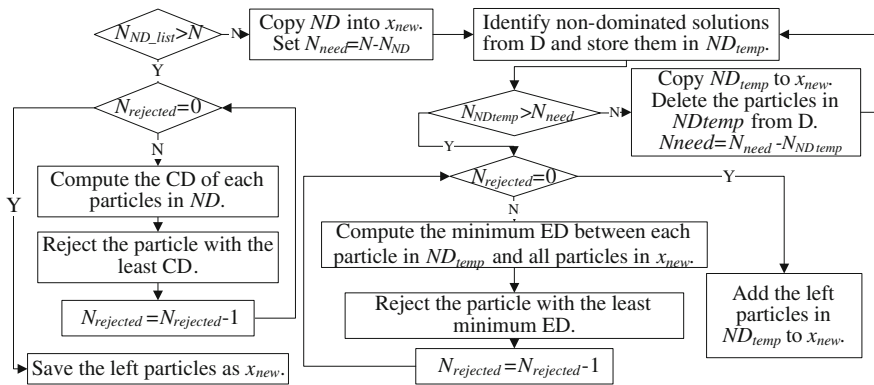
Fig. 3 Scheme of selecting the particles for next iteration

of CD based sorting technique stated above, SR is applied to the selection of particles for the next iteration. The scheme is illustrated in Fig. 3. The procedure is:

1. Identify the non-dominated solutions from R and store them in a matrix ND, while store the dominated solutions in a matrix D. Set $x_{new} = []$.
2. Reject $N_{ND} - N$ particles from ND according to SR strategy if $N_{ND} > N$, save the rest particles as x_{new} for the next iteration and end. Else if $N_{ND} < N$, copy the particles in ND to x_{new} and go to c), set $N_{need} = N - N_{ND}$.
3. Identify non-dominated solutions from the current dominated solutions in D and store them in ND_{temp}, the number of whom is N_{NDtemp}.

4. If $N_{NDtemp} > N_{need}$, set $N_{rejected} = N_{NDtemp} - N_{need}$ and execute the following operations, else go to e). Compute the Euclidian distances (EDs) between each particle in ND_{temp} and all particles in x_{new}, obtain the minimum ED for each particle in ND_{temp}. Then reject the particle with the least minimum ED. Repeat until N_{need} particles left in ND_{temp}. Add the left particles to x_{new} and end.

5. Add the particles in ND_{temp} to x_{new} and delete them from D. Let $N_{need} = N_{need} - N_{NDtemp}$.Go to c).

As stated, the non-dominated solutions at the current iteration are identified from the combination of the current particle swarm and its update. And particles for next iteration are selected from the current non-dominated solutions (If $N_{ND} < N$, some are selected from the dominated). Since the plus selection is adopted, an external archive [7] is not necessary. In other words, the relationship $N_{ND} \leq 2N$ is true at any iteration. Thus, it's different from using external archive that additional techniques need to be used to prune the archive.

2.3 MOPSO-SR Algorithm

Rapid loss of diversity within the swarm results in premature convergence. In case of MOPSO, appropriate promotion of the swarm diversity is critical to control premature convergence and enhance the global convergence to the true Pareto front. Techniques [3, 4] for maintenance of swarm diversity and enhancement of global search ability through the selection of leader particles, means of inertia weight and mutation operator are described along with the MOPSO-SR algorithm.

1. *Initialization*. Set population size N and maximum iteration M_t. Initialize the population P with velocity v and position x. Set personal best position $p_i = x$.
2. *Update leader particle and p_i for each individual*. Identify the non-dominated particles from P. For each particle, dynamically generate a set of weights, w_{t1}, w_{t2}, ..., w_{tM}, and compute fitness of each non-dominated particle according to Eq. (1) and assign the non-dominated particle with the largest *fitness* value as the individual's leader particle.

$$fitness = 1 \Big/ \sum_{i=1}^{M} w_{ti}f_i, \quad w_{ti} = \lambda_i \Big/ \sum_{i=1}^{M} \lambda_i, \quad \lambda_i = U(0,1) \tag{1}$$

where M is the number of objectives, and f_i is the ith objective. The function $U(0,1)$ generates a uniformly distributed random number within the interval [0,1].

3. *Generation of new population R*. Generate P's offspring population Q based on x and v according to Eq. (2), Combine P and Q together to get R.

$$v(t + 1) = wv(t) + c_1 r_1 [p_i - x(t)] + c_2 r_2 [p_g - x(t)]$$
$$x(t + 1) = x(t) + x(t + 1)$$
(2)

where t is the current iteration, w is inertia weight and a dynamically changing one formulated as Eq. (3) is used in the study, c_1 and c_2 are acceleration coefficients, and r_1, $r_2 \in [0,1]$ are random numbers with uniform distribution.

$$w(t) = w_0 + r_3 * (1 - w_0), \quad w_0 \in [0, 0.5]$$
(3)

where r_3 is a random number uniformly distributed in [0, 1].

4. *Identification of non-dominated particles of current iteration from R*. Store the non-dominated particles in a matrix *ND*, store the dominated ones in *D*.
5. *Selection of particles for next iteration according to* **2.2**.
6. *Mutation operation*. If needed, execute a mutation operation [4], else go to 7.
7. Return to Step 2 until M_t is met.
8. Store the non-dominated particles as the final Pareto solutions.

3 Demonstration of Effectiveness of MOPSO-SR

To demonstrate its competitiveness, MOPSO-SR is compared with several MO algorithms that are representative of the state-of-the-art, NSPSO [8], NSGA-II [2], σ-MOPSO [9] and MOPSO [10], together with AWPSO [3] and AEPSO [4]. Four commonly recognized benchmark functions ZDT1-ZDT4 are considered to be difficult because of the large number of decision variables, disconnectedness of true Pareto solutions, and multiple local fronts [7]. Two performance metrics [2], namely the generational distance GD and the spread S, which measure the closeness of the Pareto solutions identified by MOPSO-SR to the true Pareto solutions in the objective space and diversity of Pareto solutions along the Pareto front, respectively. For all test functions handled with MOPSO-SR, $w_0 = 0.3$, $V_{limit} = 0.2V_{max}$, $\beta = 0.1$, $N = 100$, $M_t = 250$. In order to establish repeatability, the proposed algorithm runs 30 times independently. Table 2 summarizes a comparison between mean values of convergence and diversity on ZDT1 $\sim$ ZDT4, performed by MOPSO-SR and other algorithms. Results indicate that the proposed algorithm performed very well as far as convergence and diversity are concerned. MOPSO-SR has achieved dramatically better convergence on all test functions and better diversity on ZDT3 and ZDT4, while σ-MOPSO has better diversity on ZDT1 and ZDT2. Figure 4 depicts the Pareto fronts identified by MOPSO-SR. As observed from Fig. 4, MOPSO-SR is capable of attaining the true Pareto front. Techniques adopted to maintain swarm diversity and to enhance global search ability have resulted in the better convergence performance.

Table 2 Mean value and var value of convergence measure and diversity

	Index	ZDT1		ZDT2		ZDT3		ZDT4	
		Mean	Var	Mean	Var	Mean	Var	Mean	Var
[8]	GD	0.00642	0.00000	0.00951	0.00000	0.00491	0.00000	4.95775	7.43601
	S	1.22979	0.00484	1.16594	0.00768	0.78992	0.00165	0.87046	0.10140
[9]	GD	0.03348	0.00476	0.07239	0.03169	0.11450	0.00794	0.51305	0.11846
	S	0.68132	0.01335	0.63922	0.00114	0.83195	0.00892	0.96194	0.00114
[2]	GD	0.01638	0.00048	0.00584	0.00000	0.10205	0.00238	3.83344	1.87129
	S	**0.39856**	0.00731	**0.38927**	0.00458	0.76016	0.00349	0.82842	0.00054
[10]	GD	0.00133	0.00000	0.00089	0.00000	0.00418	0.00000	7.37429	5.48286
	S	0.84816	0.00287	0.89292	0.00574	1.22731	0.02925	1.01136	0.00072
MOPS	GD	**0.00015**	0.00000	**0.00008**	0.00000	**0.00060**	0.00000	**0.00016**	0.00000
O-SR	S	0.58881	0.00066	0.55832	0.00051	**0.50605**	0.00023	**0.55595**	0.00103
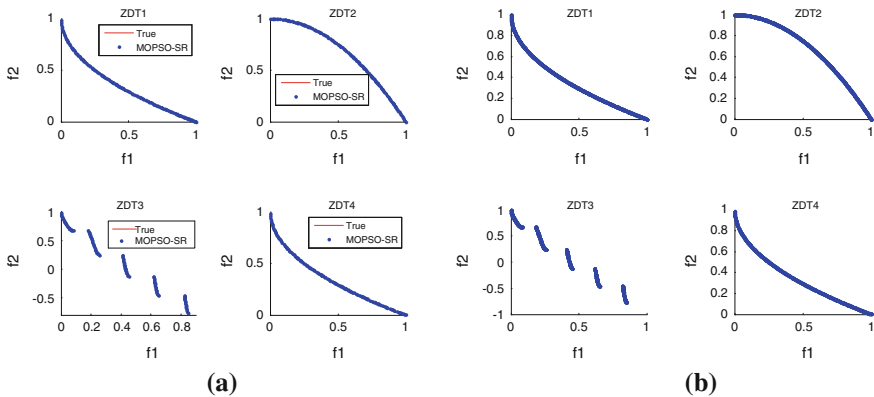
Fig. 4 Pareto fronts identified by MOPSO-SE **a** A single run, **b** ten runs

As illustrated in Table 3 and Fig. 5, compared with AWPSO and AEPSO, which employed random selection and NSGA-II CD based sorting, MOPSO-SR provided well-distributed Pareto solutions of better diversity while keeping same level convergence. Better diversity and distribution attribute to the introduction of SR technique that avoids the drawbacks of randomness caused by random selection and non-uniform spreading resulted by CD based sorting.

Table 3 Mean value of convergence measure and diversity

	Index	ZDT1		ZDT2		ZDT3		ZDT4	
		Mean	Var	Mean	Var	Mean	Var	Mean	Var
[3]	GD	0.00010	0.00000	0.00009	0.00000	0.00053	0.00000	0.00033	0.00000
	S	0.78693	0.00185	0.75830	0.00200	0.92928	0.00196	0.68290	0.00260
[4]	GD	0.00018	0.00000	0.00008	0.00000	0.00060	0.00000	0.00019	0.00000
	S	0.71169	0.00179	0.71053	0.00260	0.67291	0.00242	0.70641	0.00237
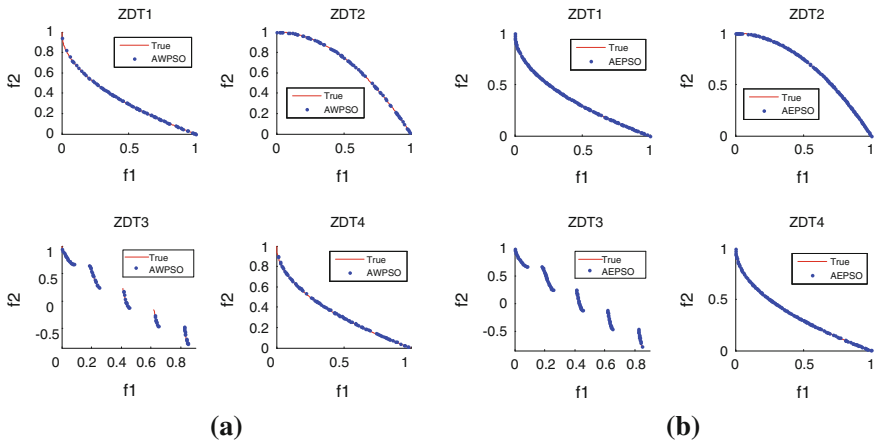
Fig. 5 Final fronts on ZDT1 ∼ ZDT2 identified by **a** AWPSO, **b** AEPSO

4 Conclusion

MOPSO-SR put forward in the study attempts to promote the diversity of Pareto solutions while keeping good convergence. This is accomplished by introduction of SR technique and techniques for maintenance of swarm diversity and enhancement of global search ability through the selection of leader particles, means of inertia weight and mutation operator. The SR technique applied to selection of particles for next iteration realizes the selection of certain solutions from a given solution set via step-by-step rejection instead of direct selection based on CD. In the comparative study, MOPSO-SR is compared against existing reprehensive MO algorithms through the use of established benchmarks and metrics. Investigations indicate that MOPSO-SR performs well considering convergence and diversity. Though, it has good stability and achieved dramatically better convergence on all test functions and better diversity on ZDT3 and ZDT4, it doesn't have the best diversity on ZDT1 and ZDT2. Hence, promotion of MOPSO-SR robustness needs further study.

Future work will also look into application of MOPSO-SR to problems with more than two objectives even with many objectives, and also real-world multi-objective optimization problems.

References

1. Zitzler E, Deb K, Thiele L (2000) Comparison of multiobjective evolutionary algorithms: empirical results. Evol Comput 8(2):173–195
2. Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput 6(2):182–197
3. Mahfouf M, Chen MY, Linkens DA (2004) Adaptive weighted particle swarm optimisation for multi-objective optimal design of alloy steels. Lect Notes Comput Sci 3242:762–771
4. Chen MY, Zhang CY, Luo CY (2009) Adaptive evolutionary multi-objective particle swarm optimization algorithm. Control Decis 24 (12):1851–1855, 1864
5. Ping H, Jin-yang Y, Yong-quan Y (2011) Improved niching multi-objective particle swarm optimization algorithm. Comput Eng 37(18):1–3
6. Zhang L, Xu Y, Wang Z, Li X, Li P (2011) Reactive power optimization for distribution system with distributed generators. Trans China Electrotechn Soc 26(3):168–173
7. Sierra MR, Coello CAC (2006) Multi-objective particle swarm optimizers: a survey of the state-of-the-art. Int J Comput Intell Res 2(3):287–308
8. Li XD (2003) A non-dominated sorting particle swarm optimizer for multiobjective optimization. Lect Notes Comput Sci 2723:27–48
9. Mostaghim S, Teich J (2003) Strategies for finding good local guides in multi-objective particle swarm optimization (MOPSO). In: Proceedings of the 2003 IEEE swarm intelligence symposium, pp 26–33
10. Coello CAC, Pulido GT, Lechuga MS (2004) Handling multiple objectives with particle swarm optimization. IEEE Trans Evol Comput 8(3):256–279

IRSR: Recover Inter-Domain Routing System from a Higher View Beyond Internet

Yu Wang, Zhenxing Wang and Liancheng Zhang

Abstract Nowadays security situation of Internet is getting surprisingly worse. Many studies show that under the intensive paralyzing attack, inter-domain routing system which acts as the critical infrastructure of Internet, will falls into large-scale and long-term failure both of the routing nodes and links, endangering the running performance of Internet. Based on the centralized control theory, IRSR model proposed here builds an independent Decision Center above the existing inter-domain routing system, it could provide global situation awareness by using the sensor networks and controller networks deployed in each AS, and implements fast recovery from failures based on methods including pre-computed failover topology and consistent view. IRSR guarantees the maximum compatibility with existing routing protocols and structure of inter-domain routing system, which will reduce the deployment cost and complexity. Moreover, it also overcomes the problems like concealed AS information and long BGP convergence, which improves the recovery velocity and coverage rate.

Keywords Inter-domain routing system · Paralyzing attack · Route recovery · Pre-computed failover topology · Consistent view

Y. Wang (✉) · Z. Wang · L. Zhang
National Digital Switching System Engineering & Technological Research Center,
Zhengzhou 450002, China
e-mail: stonchor@gmail.com

Z. Wang
e-mail: Zhenxingwang@gmail.com

L. Zhang
e-mail: A1ien@gmail.com

1 Introduction

Inter-domain routing system plays an important role in nowadays economic and social development, deeply influencing the national security [1]. However, severe event involving inter-domain routing system happens frequently and the attack methodology is turning more and more complicated which threatens the whole Internet, meanwhile, the remedy and recovery cost is getting much higher.

The reasons for inter-domain routing system failure include intended attack, random fault, misconfiguration and accidental disaster. According to the uneven connectivity distribution of Internet [2, 3], we observe that this type of structure possesses good robustness in the case of random failure. However, in terms of the intended attacks which are deliberately planned and of great intensity, Internet will be extremely fragile in that some key routers' crash could directly impact on the connectivity of large-scale networks, resulting in the network services and functions disrupted and paralyzed.

Considering the regional or national network security, studies for fast recovery from the inter-domain routing system failure under paralyzing attacks are exigently needed. On the one hand, intensity of network paralyzing attack is growing rapidly which is likely to bear the inter-domain routing system down, whereas the BGP convergence has to take a long time to work. On the other hand, within the spreading scale of Internet, many applications especially of real-time requirements ask for more survivability and stability.

Centralized control [4, 5] is to create a logically independent administrative domain, and directly control the network system. The paradigm of "directly control" means that the decision units of administrative domain will directly and specifically establish routing status on the nodes, but not configure other progress and then calculate the routing status. This paradigm is able to simplify the network control to a great extent, and improve the overall running efficiency.

To confront the potential failure of inter-domain routing system under paralyzing attacks, an inter-domain routing system recovery (IRSR) model is proposed. Different from existing mechanisms, the Border Gateway Protocol (BGP) will not be changed, and specifically the IRSR model establishes an independent decision center above the inter-domain routing system which acts as a centralized administrative domain. By virtue of the sensor network and controller network deployed in each AS, IRSR will provide the global situation awareness and implement route recovery based on methods including pre-computed failover topology and consistent view.

2 Related Works

Many different schemes for routing system recovery under the failure situation have been explored and proposed. This section introduces these methods and especially explains the relationship between IRSR model and typical existing efforts.

Consensus Routing [6] references consistent snapshot method of the distributed systems and hold a cyclically consistent view of inter-domain routing system. However, the deficiencies include (1) increasing the work load of ASes; (2) when network is almost paralyzed, information cannot be distributed to underlying ASes; (3) BGP protocol needs to be modified.

FCP [7] uses a distribution platform to release global consistent view of topology. The disadvantages lie in that: (1) routers need to recalculate paths for almost every failed message that leads to overloading of routers; (2) packets need to be modified whose length is variable and bear a big overhead.

R-BGP [8] realizes the routing system recovery based on method of pre-computed failover path, which solves the problem of route selection and distribution. However, RCN has to be added into each update message which modifies BGP semantic.

RRL [9] is able to provide a small amount of simple and intuitive network failover topology for the recovery. However, RRL can only be applied to the intra-domain routing recovery. MRC [10] made a further improvement based on RRL by limiting the link weights to obtain failover routes.

To sum up, the starting point of existing solutions can be attributed to two aspects: building a globally consistent view or using pre-computed failover topology. We propose an inter-domain routing system recovery model integrating their merits. To avoid existing deficiencies, the basic principles which must be considered when designing the model include:
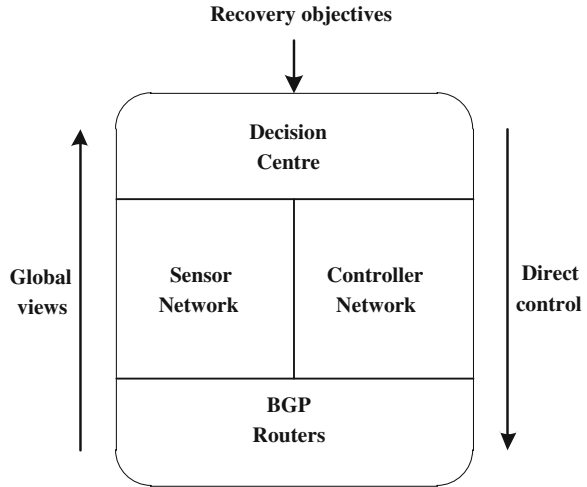
Principle 1: No modifications to any routing protocol. Principle 2: No more workload for router in respect of calculation and storage. Principle 3: Should not rely on existing network topology to distribute the recovery paths. Principle 4: Need not know the detailed mechanism of potential paralyzing attack during recovery process. Principle 5: Perform fast route recovery in multi-failure scenario.

3 IRSR Model

3.1 Architecture of IRSR Model

The key feature of IRSR model is to construct a physically independent decision center above the target inter-domain routing system to implement centralized control. When routing system is nearly or already in a paralyzed condition, the decision center is able to avoid the dependence of underlying channel to transmit control messages, utilizing the knowledge of global topology and running status of the routing system, choose appropriate pre-computed failover topology and directly reconfigure the routing tables and forwarding tables of AS routers in a synchronous way, so as to quickly complete the routing system recovery.

For inter-domain routing system of Internet, IRSR model is established using the access control relationships and the abstractive framework is shown in Fig. 1.

Definition 1 ISIR is able to monitor and recover multi-AS which are adjacent and associated. For the given AS_a, $R_a = \{r_{a1}, r_{a2}, \ldots, r_{ak}\}$, where R_a denotes the set of BGP routers in AS_a, and r_{ai} represents the i-th router within the set.

Definition 2 Let DC (Decision Center) acts as the top-level decision-making center of the IRSR model. DC depends on SN (Sensor Network) and CN (Controller Network) deployed in ASes to implement information collection and failure recovery for inter-domain routing system.

Definition 3 Suppose $SN = \{SN_a, SN_b, \ldots\}$ and $CN = \{CN_a, CN_b, \ldots\}$. For $SN_a = \{s_{a1}, s_{a2}, \ldots, s_{an}\}$, SN_a denotes the sensor network deployed in AS_a, and S_{ai} represents the i-th sensor of SN_a set. Similarly, for $CN_a = \{c_{a1}, c_{a2}, \ldots, c_{am}\}$, CN_a denotes the controller network deployed in AS_a, and c_{ai} represents the i-th controller of CN_a set.

Inter-domain routing system protected by IRSR consists of multiple ASes. To maximize the protection of AS privacy, for a single AS_a, it is only the sensor network SN_a and controller network CN_a have the right to directly operate the BGP routers of AS_a. In addition, AS can flexibly participate in or withdraw from this recovery architecture which improve the scalability and flexibility of this model.

Definition 4 IRSR model includes two relationships, i.e., router resource access and route control.

We let $A = \{(s, r) \in SN \times R : \text{Sensor } s \text{ can get access to the resources of router } r\}$, which defines the relationship between the sensor network and target router resources indicating sensor node s is able to access the resources of the router r; And also, $B = \{(c, r) \in CN \times R : \text{controller node } c \text{ can control router } r\}$
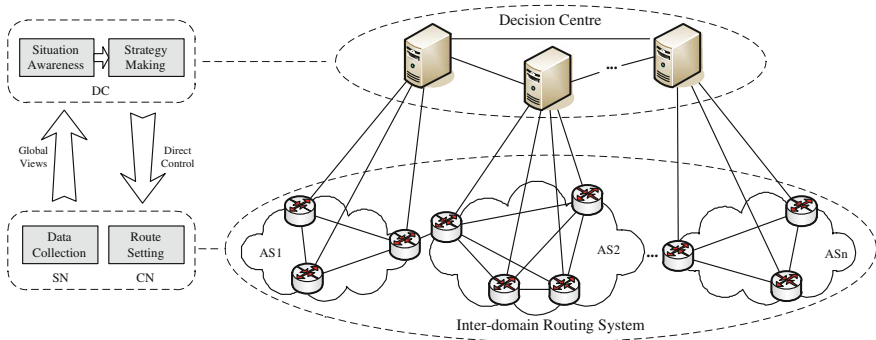
Fig. 2 Illustration of IRSR deployment at AS level

defines the control relationship between control network and the target router, indicating sensor node c is able to control the resources of the router r.

Figure 2 shows the abstractive AS-level deployment of IRSR. It is observed that IRSR works on two layers: (1) sensor network and controller network are deployed in each AS, respectively playing the role of status collection and route configuration; (2) with their help, decision center beyond them could receive the global view for situation awareness, and then make recovery strategies to control the BGP routers of each AS directly through the controller network.

Specifically, sensor network is responsible to collect various types of resource information of inter-domain routing system, in order to identify potential anomalies which may indicate the decline of network service quality or failure events because of attacks. The main collection of information include: processor utilization, buffer usage, ingress/egress traffic volume and their ratio, number of input/output errors, number and frequency of link failures, as well as number of authentication failures. After that, the sensor network should aggregate all the information collected by sensors and organize them into encrypted data stream based on time series. Finally, these content are presented up to the decision center.

The decision center is in charge of the generation of control commands. First, data stream from ASes is aggregated which is used to determine the potential anomalies by comparing the newly collected network resource information with existing historical behavior record. Then, DC utilizes the feature detection to recognize anomalies and the security situation can be evaluated, based on which, the global view of inter-domain routing system security situation can be acquired. At last, according to the different real-time status, DC makes control strategies and distributes commands through the corresponding controller networks.

The controller network implements control operations to the specific AS. After receiving the data stream from DC, the commands are parsed and distributed into each controller. According to differences of the recovery strategies, the control progress may include: using the correct topology structure in real-time status to reconfigure the routing table and forwarding table of BGP routers synchronously, which could avoid the time-consuming convergence process. Startup the backup

paths or routers when the infrastructure of inter-domain routing system suffered the physical destruction and this strategy could quickly resume normal routing function and ensure the network service without interruption.

## 3.2 FSM of Inter-Domain Routing System with IRSR

Definition 5 S is used to describe the states of different moments of the inter-domain routing system, and $S_i(t)$ represents the state at time t.

In order to dynamically depict the state transition of inter-domain routing system under IRSR, S is divided into three categories: the initial state where $S = 0$; paralyzed state where $S = -1$; frozen state where $S = 1$. Figure 3 shows the responding finite state machine of inter-domain routing system under IRSR model.

Initial state ($S = 0$) indicates the normal state of routing system that it has not yet suffered any attack or the system has just been recovered. Paralyzed state ($S = -1$) means the inter-domain routing system has been attacked where large-scale communication services almost fail or have been broken down. At that time, decision center is aware of the paralyzing situation through the sensor network, and then the inter-domain routing system is turned into the frozen state ($S = 1$) by IRSR during which time the packets processing and BGP updating of routers are stopped in a compelling way for a short period.

With the help of sensor network, IRSR is always monitoring the operational status of the inter-domain routing system. When severely dangerous situations inside AS_a such as the router processing lapses sharply or masses of control messages emerges, are sensed by the decision center, thus it is necessary to bring AS_a into the frozen state. The purpose is to make AS_a in a temporary isolation so that to avoid further and heavier load for the routing system. At the meantime, routing information of other ASes which connect with AS_a has to be adjusted

Fig. 3 FSM of Inter-domain routing system under IRSR model
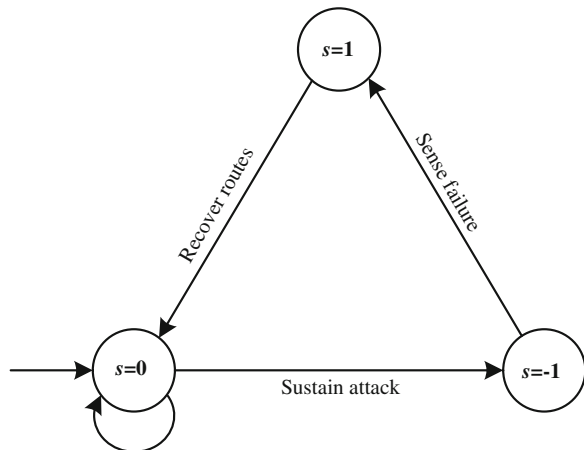
correspondingly. Then, the source of attacks should be determined through the analysis of status information represented by CN_a, and the recovery strategy will be adopted (this part will be expatiated in Sect. 5). AS_a is not allowed to rejoin inter-domain routing system until the communication has been in the normal state, after that, routes information of neighboring ASes will be reconfigured synchronously.

When the entire inter-domain routing system has been attacked by the sudden assault and multi-AS has been in the paralyzed situation, IRSR will mandate it into the frozen state. Decision center will analyze the key paralyzed nodes or links as soon as possible, and start the backup infrastructure. The pre-computed failover topology based on the very situation is utilized to reconfigure BGP routers so that to recover the inter-domain routing system into the initial state.

4 Discussions

4.1 Advantages of IRSR

Compared with existing recovery approaches in the current security situation for inter-domain routing system, advantages of IRSR are reflected in:

1. Focusing on the confrontation to potential high-risk paralyzing attack.
   Once the inter-domain routing system gets paralyzed under attack, the whole network service will be stagnated. Existing attacks towards inter-domain routing system are not only harmful, but also difficult to confront with current defense architecture. Therefore, the research of IRSR model is of strong relevance and necessity for the inter-domain routing system security.
2. Maintain existing protocols and structure of inter-domain routing system.
   Taking into account the current ASes deployment scale and complexity, clean-slate design and modifications to the existing routing protocols and infrastructure are not feasible neither. It is the decision center which is established independently above the inter-domain routing system, to perform global monitoring and recovery through the underlying sensor network and controller network.
3. Avoid the BGP convergence and complete the fast recovery.
   When inter-domain routing system is in danger for the paralyzing attacks, the decision center forces it into frozen state, and begins to implement the recovery operation. Since the recovery process needs no interactions with neighboring routers, the phenomenon of update delay and inconsistent view caused by BGP convergence can be avoided completely.
4. High flexibility and scalability.
   IRSR belongs to the centralized infrastructure, in which the decision center is able to make overall strategies. The key algorithms for situation awareness, as well as the control strategies are not subject to the number of ASes and their adjacency relationships. Administrative ISP of each AS has the right to decide whether entering or leaving this defense model according to its own security needs.

4.2 Challenges for IRSR

It is of great significance for IRSR to confront paralyzing attacks inter-domain routing system, and the advantages are distinct and promising. However, design and implementation of IRSR must face challenges that mainly include three aspects.

1. Collaborations of multi-AS in the inter-domain routing system.
   The biggest challenge for IRSR model is how to build a collaborative relationship between ASes. AS usually hides privacy information which impedes the information sharing among ASes. However, network administrator should weigh the pros and cons between the needs of state infrastructure security and ISPs' privacy protection. As we recommend, vulnerability and security need of today's inter-domain routing system should be valued.
2. Global situation assessment for inter-domain routing system.
   Nowadays we could not fully realize the specific manner and time about the attack. It is necessary to implement real-time global situation assessment for inter-domain routing system. Main steps include: (1) perceptions for the exception information in routing system; (2) establishment of assessment algorithm based on the multi-source information fusion; (3) using assessment algorithm to evaluate the security status of single router, AS and the entire routing system.
3. Communication efficiency and security between DC and underlying networks.
   Bandwidth, traffic volume, communication protocols and other factors between SN and CN will influence the practical efficiency and security of IRSR. Therefore, we should focus on the factors involving the data and frequency during the sensors collect, and reduce the data volume as much as possible.

5 Conclusions

Inspired by the existing methods of consistent view and pre-computed failover topology, this paper proposes an inter-domain routing system recovery model based on the centralized control mode. Without changing the basic routing protocols of nowadays Internet, IRSR adopts the cross-layer and mandatory recovery solutions, utilizing the independent decision center to monitor and control the inter-domain routing system, and implement fast recovery for the paralyzed routing system. The IRSR model depicted here is only a rough skeleton, and future works include reinforcement of the recovery model architecture, incentive mechanism for collaborations between ASes, as well as the cooperative control mechanism for the controller network.

References

1. Lippert KJ, Seger RA (2011) The story behind network defense. In: Proceedings of the 8th international conference on information technology: new generations. IEEE society, pp 909–913
2. Carlson JM, Doyle J (2002) Complexity and robustness. Proc Nat Acad Sci 99(Suppl 1):2538–2545
3. Bollobas B, Riordan O (2004) Robustness and Vulnerability of scale-free random graphs. Int Math 1(1):1–35
4. Iqbal H, Znati T (2010) Overcoming failures: fault-tolerance and logical centralization in clean-slate network management. In: Proceedings of IEEE INFOCOM 2010
5. Yan H (2010) A practical system for centralized network control. Carnegie Mellon University
6. John JP, Katz-Bassett E, Krishnamurthy A (2008) Consensus routing: the internet as a distributed system. In: Proceedings of the NSDI 2008, USENIX association, pp 351–364
7. Lakshminarayanan K, Caesar M, Rangan M (2007) Achieving convergence-free routing using failure-carrying packets. In: Proceeding of the ACM SIGCOMM 2007. ACM Press, Kyoto, pp 241–252
8. Kushman N, Kandula S, Katabi D (2007) R-BGP: Staying connected in a connected world. In: Proceedings of the NSDI 2007, USENIX association, pp 341–354
9. Hansen AF, Kvalbein A, Cicic T (2005) Resilient routing layers for recovery in packet networks. In: Proceedings of the dependable systems and networks (DSN), IEEE computer society, pp 238–247
10. Kvalbein A, Hansen AF, Cicic T (2006) Fast IP network recovery using multiple routing configurations. In: Proceedings of the INFOCOM 2006, IEEE computer society, pp 1–11

A New Parallel Genetic Algorithm Based on TriBA Topological Structure

Kang Sun and Wook Ahn Chang

Abstract In order to advance the speed of solving a large combinatorial problem, this paper proposes a new master-slave parallel genetic algorithm (PGA) based on the triplet based architecture (TriBA) topological structure. With the special topological structure by which the problem can be mapped into this model, the TriBA is able to realize the parallel computing in child topological structures and the parallel real-time communication. The theoretical analysis proves that the proposed TriBA PGA can enhance the computation speed and decrease the communication costs, thereby resulting in a novel PGA model to handle the large combinatorial problems.

Keywords TriBA topological structure · Parallel genetic algorithm · Master-slave model · Computation speed · Communication costs

1 Introduction

Genetic algorithm (GA) is a global random search strategy [1, 2]. When solving a combinatorial problem, GA is efficient and robust. With the growth of the problem's scale and the development of the multi-core computer, parallel genetic algorithm (PGA) has been widely used in many fields, such as image processing,

K. Sun (✉) · W. A. Chang
Department of Computer Engineering, Sungkyunkwan University (SKKU),
2066 Seobu-ro, Suwon 440-746, Republic of Korea
e-mail: sunkangskku@gmail.com

W. A. Chang
e-mail: cwan@skku.edu

pattern recognition, and so on [2–4]. As the problem's scale increases, the parallel computing speedup and the communication efficiency have been enhanced rapidly [2, 4, 5]. This paper presents a new parallel genetic algorithm model that employs the Triplet Based Architecture (TriBA) topological structure model [6]. Using the features of the TriBA's special structure and the computing mode, it can solve the combinatorial problem within the multi-level parallel computing framework and decrease the communication time among the nodes.

2 TriBA's Features

2.1 TriBA Topological Structure

Triplet Based Architecture (TriBA) is a new computer topological structure of multi-core on chip. It is an expanded version of the all-connected topologies; every node in ground-grade is connected with other three nodes around. A single node is connected with three communication links, like a triplet. When iterating this process K times, the number of nodes becomes 3^K. Figure 1 demonstrates a set of TriBA topological structures. When $K = 2$, for instance, the original population can be divided into three sub-populations in accordance with the tree triplets; it looks like three TriBA structures with $K = 1$. Thus, the TriBA's topology can enhance the speed of computation. With respect to the population of PGA, this structure can be easily expanded by increasing the number of cores (i.e., nodes).

2.2 TriBA's Computing Mode

The TriBA's topology retains four basic computing modes: single computing, center computing, parallel computing, and level computing. As shown in Fig. 2, these special computing modes would provide a better computing environment for the parallel computing (i.e., PGA).
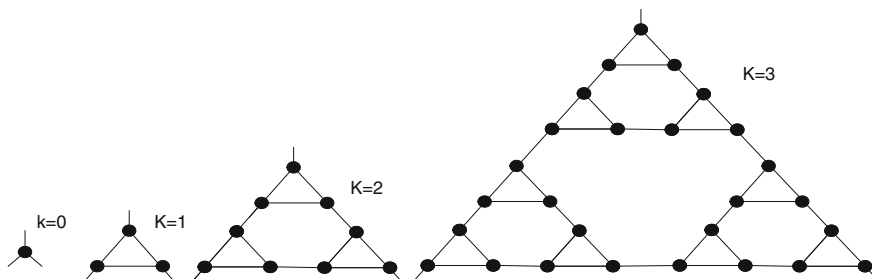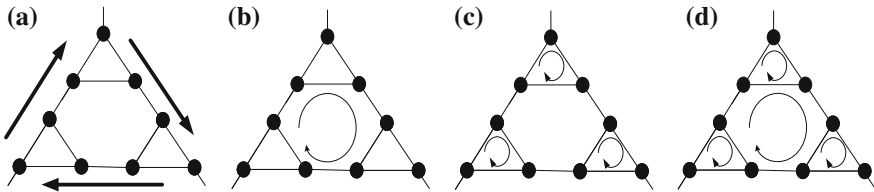

Fig. 1 TriBA's structures

Fig. 2 TriBA's computing modes. (**a**) Single Computing. (**b**) Center Computing. (**c**) Parallel Computing. (**d**) Level Computing

3 Proposed Approach: TriBA PGA

3.1 Overall Design

As the traditional parallel GA's model, the parallel computing is always used in the process of computing suitability. In the process of migration and mutation, they always exploit serial computing. More time wastes in these processes and communications among nodes are required. Using the TriBA's structure, TriBA parallel genetic algorithm can be classified as an island parallel genetic algorithm with the TriBA communication topology. Because of the new computing mode and topological structure, it is fitter in parallel computing than that in the traditional model [2, 7, 8].

3.2 Operational Procedures (K = 2)

1. **Initialization**

   The main process gets the parameters for the original population, such as demes, the maximum generation, mutation rate, and crossover rate. The original population is divided into three sub-populations.

2. **Parallel computing in sub-processes**
   (a) Get the parameters from the broadcasting in the main process.
   (b) Using the parallel computing mode, perform the evolutionary operation such as selection, crossover, mutation, and fitness evaluation.
   (c) Using the center computing mode, compare the best individuals of the different sub-populations. If the new best is better than the old best, store the new one through the center mode.

3. **Output**

   When the number of generations reaches the maximum generation or the desired quality is achieved, the suitable individuals would be in the process of center computing. Put these individuals to the main process, output them.

4 Speedup Analysis of TriBA PGA

The analysis of execution time focuses on the control processor in the TriBA topological structure and the master processor in the traditional master-slave model. T_c is the communication time and T_f is the computing time. Considering all the contributions from the communication and the computation, the total elapsed time for one generation of the master-slave model is given as [4]

$$T_{P1} = PT_C + \frac{nT_f}{P} \tag{1}$$

The TriBA model can perform the computing in a single generation because data can be transferred in parallel due to its triangle structure. Thus, the total time for the TriBA model is given as

$$T_{P2} = 2T_c \log_3 P + \frac{nT_f}{P} \tag{2}$$

When more slaves use, the computation time decreases as desired, but the communication time increases. To find the optimal number of processors, let $dT_p/dp = 0$ and thus, we have

$$P_1^* = \sqrt{\frac{nT_f}{T_c}} \tag{3}$$

$$P_2^* = \frac{nT_f}{2T_c} \tag{4}$$

Let Eqs. (3) and (4) be equal, we have $P_1^2 = 2P_2$ since $P_2 \geq 3$. Thus, the TriBA model would be more effective than the traditional master–slave model. Then, the speedup of TriBA PGA is compared with that of the master-slave PGA. This can be computed by Eq. (5), where p is the number of processes and T denotes T_c/nT_f. The speedup is related to the communication time (cost), the number of processors, and the original population size. Since p is 3, 9, and so on, Eq. (5) is greater than '1.0'. Therefore, the TriBA's topology is better than the traditional master-slave's one.

$$\frac{T_{P_1}}{T_{P_2}} = \frac{p^2 T_C + nT_f}{2p^2 T_c \log_3 p + nT_f} = \frac{pT(p - 2\log_3 p)}{2pT \log_3 p + 1} + 1 \tag{5}$$

5 Experimental Results

The validity of the proposed approach is verified by the computational experiments. Figures 3 and 4 are the optimization curves when the number of nodes is 3 and 9, respectively. With the increasing of population's number and the change of

Fig. 3 TriBA's and master-slave's optimization curve with 3 nodes
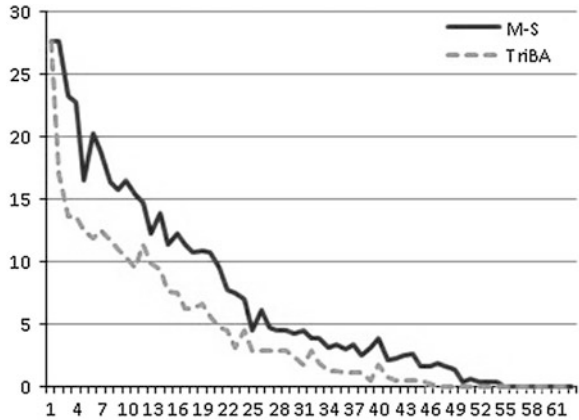
Fig. 4 TriBA's and master-slave's optimization curve with 9 nodes
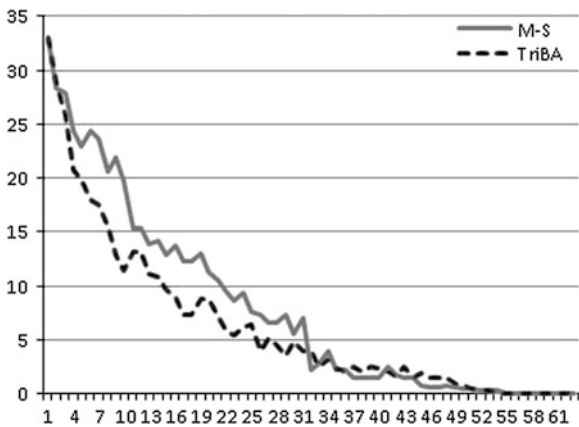
topological structure, the performance was enhanced at every side. Therefore, the proposed TriBA PGA gets a better computing efficiency than the traditional one.

The speedup of 3 nodes and 9 nodes is exhibited in Table 1. It is confirmed that the proposed algorithm is more efficient than the traditional master-slave PGA. Even though the test of TriBA's computing more than 9 cores has not been conducted, we can conclude that the new topological structure is able to get a better speedup in the multi-core environment.

Table 1 Speedup comparison in TriBA and M-S structures

	3 nodes	9 nodes
TriBA generations	52	55
M-S generations	55	61
Speedup	1.057	1.10

Table 2 Estimate speedup on 27 cores and 81 cores

No. of cores	3	9	27	81
Speedup	1.057	1.10	1.875	3.805

Considering the speedup of 3-processor and 9-processor TriBA topologies, it can be turned out that the ratio of the computing time T_C and the communication time T_f is 0.0556. Using Eq. (5), the estimate speedup for 27 cores and 81 cores can be carried out. The results are demonstrated in Table 2. Note that this improvement can be achieved by employing the TriBA structure.

6 Conclusion

This paper has addressed the TriBA-structure model as a promising topological structure of PGA. It combines the TriBA's topology with the traditional master-slave PGA. Using the feature of TriBA, we could reduce the cost and delay of communication. Comparing with the traditional master-slave model, the proposed model achieved a better speedup performance. Future work includes more extensive experiments to get more insight behind this TriBA model. Also, it needs to develop a method to dynamically change the number of cores for the given population.

References

1. Manderick B, Spiessens P (1989) Fine-grained parallel genetic algorithms. In Proceeding of the third international conference in genetic algorithms, pp 428–433
2. Goldberg DE (1999) Sizing population for serial and parallel genetic algorithms. In: Proceedings of 3rd ICGA, Morgan Kaufmann
3. Nakano R, Davidor Y, YamadaBruce T (1994) Optimal population size under constant computation cost. In: Proceedings of Parallel problem solving from nature—PPSN, pp 130–138
4. Cantu-Paz E (2000) Efficient and accurate parallel genetic algorithm. Kluwer, Norwell, MA
5. Maresky J, Davidor Y, Gitler D, Barak A (1995) Selectively destructive re-start. In: Proceedings of 6th ICGA, Morgan Kaufmann
6. Cai-xia L, Feng S, Bao-jun, Q (2010) TriBA interconnection topology structure and its performance analysis. Comput Eng 36(15):doi:10.3969/j.issn.1000-3428.2010.15.037
7. Li W, Huang Y (2012) A distributed parallel genetic algorithm oriented adaptive migration strategy. In: Proceedings of eighth international conference on natural computation (ICNC), pp 592–595. doi:10.1109/ICNC.2012.6234584
8. Zhu-rong, W., Tao, J., Du-wu, C., Xin-hong, H.: A study of hybrid parallel genetic algorithm model. In: Proceedings of seventh international conference on natural computation (ICNC), pp 1038-1042. doi:10.1109/ICNC.2011.6022186

Based on Grey Linguistic Multi-Criteria Decision Making Method for Real Estate Websites' Evaluation

ZhiFeng Li and LiYi Zhang

Abstract The evaluation system of real estate website has been studied and it shows that there are still many drawbacks that lead to inaccurate evaluation information and affect the decision of house buyers. Therefore, the thesis employs the Grey Linguistic Multi-Criteria Decision-Making (GLMCDM) Model to improve the evaluation system of real estate website. By establishing the Grey Linguistic Multi-Criteria Decision-Making (GLMCDM) Model, customers' evaluation information is analyzed, then the proper apartments can be recommended to house buyers on the basis of their information. Finally, the feasibility of the model is proved by calculation.

1 Introduction

Website evaluation appears along with the birth of website. Initially, the evaluation is only limited to a general assessment of the website. With the improvement of evaluation index and the development of customer relationship, evaluations concerning website product, service and after-sale service begin to appear. In the early stage, qualitative analysis was mostly used when the evaluation came into

Z. Li (✉) · L. Zhang
Wuhan University, Wuhan, China
e-mail: 593645880@qq.com

L. Zhang
e-mail: toyota99@21cn.com

Z. Li
Guangdong peizheng college, Guangdong, China

being. Later, in order to obtain more accurate data, qualitative index was given certain numerical value. That is the reason why most of current evaluations employ a method that combines qualitative and quantitative analyses together. All theses developments have contributed a lot to the perfection of website evaluation to a certain degree.

However, the current website evaluation is still far from accurate and unable to provide any referential information. For example, when purchasing, the customer finds it hard to judge the evaluation information because the information analysis is long and subjective. In addition, in light of analysis and study on previous relevant literatures, few people have studied on website evaluation. As GLMCDM Method has been deeply studied in multiple areas, in this thesis, the author attempts to apply it in analyzing website evaluation on the basis of real estate website and matching commodity house recommended to the house buyer according to his or her information.

2 Methodology

2.1 Grey Linguistic Multi-Criteria Decision-Making

Decision always exist in the life, due to the uncertainty of the information makes it difficult to make accurate choices. Therefore, in order to better address the decision-making problems, scholars have conducted research on the gray language multi-criteria decision making method, and is widely used in various fields. The evaluation of conception design is the focus and difficult point in industrial design, and the quality of the design results is directly influenced by this process. In this paper, we investigate the multiple attribute decision making (MADM) problems for evaluating the industrial design with interval grey linguistic variables [1]. To explore the issue of foreign trade competitiveness, Liu, Xuehua used interval grey linguistic variables for evaluating [2]. Above thesis all utilize the interval grey linguistic weighted geometric (IGLWG) operator to aggregate the interval grey linguistic variables corresponding to each alternative and get the overall value of the alternatives, then rank the alternatives and select the most desirable one(s). Finally, an illustrative example is given [1, 2]. In addition to some scholars directly study the gray linguistic multi-criteria decision-making method, some studies from the side to delve into. Tian, Yanfeng, Zhang, Quan and Liu, Yang had focuses on the multiple attribute decision making problems with the attribute values being linguistic evaluations [3]. Zhang, Yi and Wei, FaJie took a deep look into the concepts of gray linguistic set and gray linguistic numbers. Aiming at solving multiple criteria decision making (MCDM) problems in which the criteria weights are unknown [4]. For the study of the solution of multi-criteria group decision, also involved in the gray linguistic. A method based on the interval grey linguistic variables weighted geometric aggregation operator is presented to solve the

multiple attribute group decision making problems, in which the attribute values and the weights take the form of the interval grey linguistic variables (IGLV) [5]. Supplier selection is a multiple-attribute decision-making (MADM) problem. Since the decision makers (DMs) such as preferences on alternatives or on the attributes of suppliers are often uncertain, supplier selection becomes more difficult. So Daisuke Yamaguchi and Masatake Nagaipropose a new grey-based approach to deal with the supplier selection problem [6].

Above thesis, gray linguistic multi-criteria decision-making method is not applied to the evaluation of real estate websites. In addition, although in the end of the thesis gives examples to illustrate the feasibility, but did not clearly reflect the validity of the method uses.

3 Real Estate Website Evaluation System Existing Problems

Currently, evaluation systems of real estate website are relatively simple. Most of them have only three levels which are not clear-cut. Therefore, as everyone has different evaluation criteria, the evaluation results are bound to differ. For evaluations of the same grade, customers may express a different degree. Simple evaluation has weak referential value and fails to influence decision making. Most of real estate websites have only three levels of evaluation: good, modest and bad. Though such kind of evaluation expresses customer perception, the degree of each level is not shown in details. For example, the level "good" only roughly shows that the customer is satisfactory, but it fails to fully show the degree of satisfaction. Thus, it is not known if the customer thinks it is excellent, good or preferable. It follows that information from such a vague evaluation result is not accurate and grey nature of the evaluation influences immensely over the ultimate result.

4 Real Estate Website Evaluation Based on the Gray Linguistic Multi-Criteria Decision Making Method to Solve Problem

GLMCDM Method is used to improve the evaluation system of real estate website, which enables customers to obtain detailed and real reference information when selecting houses. Every website evaluation has its grey nature, that is, the meaning expressed is not limited to literal understanding, but has its deeper extended meaning. The combining of extended meanings behind every evaluation forms an interval. These can be viewed as grey language. Use of GLMCDM Method to improve website evaluation system enables customers to make more specific evaluations and customers who view evaluations to obtain more accurate for making more proper choices (Fig. 1).
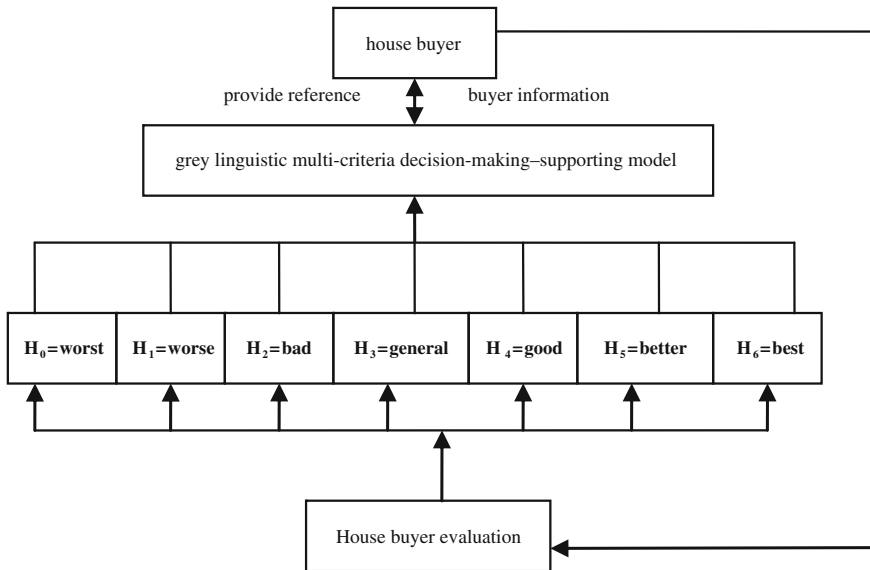
Fig. 1 grey linguistic classification

Therefore, use of GLMCDM Method to improve the evaluation system of real estate website plays an important role in solving grey and fuzzy nature of evaluation. With analyses and studies on current real estate website and combining with GLMCDM Method, the author proposes an evaluation method in which a customer grades to show degree of satisfaction. When evaluating, the customer grades on a continuum from -3 to $+3$ on the basis of satisfaction degree. -3 means terrible, while $+3$ means excellent. From the grade, we can roughly understand the customer preference. If the grade shows a positive result, we can know how good it is. Otherwise, we can also know how terrible it is. Moreover, by grading, we will have an idea about the general situation by calculating the average value. Therefore, other visitors are not easily influenced by subjective evaluation from any specific customer. Such a grading-based evaluation method is more scientific, and its results are more referential for other visitors and better at guiding in making right choices for purchasing a house.

According to the above evaluation method, the evaluation level is set to "Worst, Worse, Bad, General, Good, Better, Best" at 7 levels, and for each level with the following assumptions:

H_0 = Worst, Far below the expectations of the value of the customer perception of the product or service; H_1 = Worse, Product or service is not reach the expected effect; H_2 = Bad, Products or services substantially less than the desired effect, but there are still satisfied with the place; H_3 = General, Products or services overall to achieve the desired results, but there are not satisfied with the place; H_4 = Good, Product or service is reach the expected effect; H_5 = Better,

Customer perception of product or service value consistent with expectation; H_6 = Best, Customer perceived value of product or service far exceeds expectations.

Let $H = \{h_i | i = 1, 2, \cdots, t\}$ be a linguistic term set with odd cardinality. Any h_i represents a possible value for a linguistic variable, and it should satisfy the following characteristics, (1) The set is ordered: $h_i > h_j$, if $i > j$; (2) There is the negation operator: $neg(h_i) = h_j$ such that $j = t + 1 - i$. For example, H can be defined as

$$H = \left\{ \begin{array}{l} h_0 = \text{worst}, h_1 = \text{worse}, h_2 = \text{bad}, h_3 = \text{general}, \\ h_4 = \text{good}, h_5 = \text{better}, h_6 = \text{best} \end{array} \right\}$$

According to grey linguistic formula:

Houses	H_1	H_2	H_3
G_1	$(h_3, [0.3, 0.4])$	$(h_2, [0.3, 0.4])$	$(h_4, [0.6, 0.7])$
G_2	$(h_4, [0.4, 0.6])$	$(h_5, [0.5, 0.7])$	$(h_2, [0.4, 0.8])$
G_3	$(h_2, [0.5, 0.6])$	$(h_4, [0.5, 0.8])$	$(h_4, [0.6, 0.7])$
G_4	$(s_4, [0.4, 0.6])$	$(s_5, [0.4, 0.5])$	$(s_5, [0.6, 0.9])$
G_5	$(s_5, [0.4, 0.5])$	$(s_3, [0.5, 0.6])$	$(s_3, [0.5, 0.7])$

Step 1. Utilize the decision information given in the interval grey linguistic decision matrix $\tilde{R}$, and the GLWAA operator to derive the overall interval grey linguistic variables $\tilde{r}_i$ of the houses H_i (Let $\omega = (0.30, 0.15, 0.10, 0.30, 0.15)^T$)

$z_1 = (s_{3.2}, [0.3, 0.6])$; $z_2 = (s_{4.1}, [0.4, 0.7]5)$; $z_3 = (s_{5.6}, [0.4, 0.6])$

Step 2. We rank the above the interval grey linguistic variables by using the method presented in grey linguistic:

$$Q(z_1) = s_{3.67}, \ Q(z_2) = s_{6.12}, \ Q(z_3) = s_{5.47}$$

Step 3. The ranking of the alternatives can be gained: $H_2 \succ H_3 \succ H_1$, H_2 is the best houses for customer.

Real estate website evaluation not only provides people who have already bought houses with a channel to show their satisfaction, but also provide reference for other house buyers. Combining these two functions can give full play to the role of real estate website evaluation. Firstly, every customer makes an overall evaluation on the basis of real condition of house. Next, the customer chooses the proper evaluation grade according to his or her satisfaction degree. Then, grey linguistic multi-criteria decision-making–supporting (GLMCDMS) method analyzes and generalizes each evaluation to yield the ultimate evaluation result. Meanwhile, in accordance with house buyer information, a suited house is chosen

for the buyer. Finally, the chosen house is recommended to the house buyer. When he or she successfully buys the house, the buyer will evaluate on the basis of real condition, which will provide reference for future house buyers.

5 Conclusion

With the coming of information age and development of e-commerce, the networking of real estate is a general trend. To look for houses via internet is undoubtedly more convenient and rapid. However, rhetorical house description hardly enables the buyer to make reassuring choices, so purchasing recording and evaluation information make good reference. In this thesis, relevant analyses and studies show that current evaluation system of real estate website is not sound for it is still greatly secluded. Due to over-simple evaluation grade and fuzzy evaluation grade, those who have already purchased a house fail to make accurate evaluation and those who intend to buy a house cannot get useful reference from the evaluation. In view of this, the author employs GLMCDM Method to improve the evaluation system of real estate website. In this thesis, an evaluation method in which the customer grades within the specified scale to show satisfaction degree is proposed. GLMCDMS Method is made to show its feasibility. By making the model and calculation, the GLMCDM Method is proved to be feasible. Though this method can solve grey evaluation theoretically, yet it is not proved in practice so it is not sufficiently persuasive. Therefore, we hope that GLMCDM Method can be used in practice and applied to all aspects.

References

1. Zheng Z, Zhang J (2012) Model for evaluating the industrial design with interval grey linguistic variables. Int J Digit Content Technol Appl 6(15):136–142
2. Liu X (2011) An approach to evaluating the foreign trade competitiveness with interval grey linguistic variables. Int J Digit Content Technol Appl 5(12):51–57
3. Tian Y, Zhang Q, Liu Y (2009) Grey-relation based approach to multiple attribute decision making with linguistic evaluation. International Symposium on Intelligent Ubiquitous Computing and Education. Processing International Conference on Industrial Engineering and Engineering Management 577–580
4. Zhang Y, Wei F (2011) Gray relative analysis method for group decision making based on gray linguistic set. Processing International Conference on Industrial Engineering and Engineering Management 1:133–135
5. Liu P, Zhang X (2011) Multi-attribute group decision making method based on interval grey linguistic variables weighted geometric aggregation operator. Int J Digit Content Technol Appl 26(5):743–747
6. Daisuke Y, Masatake N (2007) A grey-based decision-making approach to the supplier selection problem. Math Computer Model 46(3–4):573–581

Generated Fast of the Aircraft Carrier Dimensions in Preliminary Design

Yu-juan Wang, Sheng Huang, Xin-yue Fang and Gang Wang

Abstract This paper studies the overall elements of aircraft carrier at the stage of demonstration and preliminary design. Firstly, the estimation formula of the attributes is developed, which should be considered in preliminary design of ship. Secondly, the basic multi-objects optimize model is built up according to weighted vector obtained by lowest deviation method. Finally, optimization based on the simulated annealing algorithm is achieved. The computing system will find the most reasonable scheme of overall elements through the model after the designer determines the constraints of the principal dimensions and the number of carrier-based aircraft and the importance of each objects.

Keywords Aircraft carrier · Multi-objective optimization · Simulated annealing algorithm · Optimization

1 Introduction

At the stage of warship demonstration and preliminary design, much schemes are needed comparative analysis and decision-making and as a foundation of the follow-up design, the scheme usually consists of a set of overall elements (principal dimensions, main hull coefficient, warship main parameters). Firstly, the

Y. Wang (✉) · S. Huang · X. Fang
College of Ship-building Engineering, Harbin Engineering University,
Harbin 150001, China
e-mail: yj020111@163.com

X. Fang
e-mail: 83172711@163.com

G. Wang
China Petroleum Liaohe Equipment Company, Panjin 124000, China

designer sets a feasible range of overall elements to compose a selection space of the elements. Then all aspects of comprehensive performance of the warship are expressed to a function of every overall element. Finally, by the employment of the single-and multi-objective optimization algorithm [1], the comprehensive performance of the warship (navigation performance, combat effectiveness, economical efficiency etc.) is taken as the target optimization and selected within the selection space by adjusting every function weight to form multiple optimal schemes for a comparative analysis and decision-making of the designer [2].

This paper sets a certain deviation range to the principal dimensions (including L, B, T, D, Δ) and the number of carrier-based aircraft of "Kitty Hawk" aircraft carrier, considering stability, the speed and powering performance, the seakeeping performance, the operational effectiveness and the cost of aircraft carrier as the target optimization. It is carried out within the deviation range by taking advantage of simulated annealing algorithm. Optimum value of overall elements are obtained to improve the warship in the five aspects of its comprehensive performance [3].

2 The Basic Principle of Simulated Annealing Algorithm

The simulated annealing algorithm is a kind of random search algorithm, and its basic thought is comparing the solution process of a certain kind of optimization with the thermal balance of statistical thermodynamics, which means to find a global optimal solution or an approximate global optimal solution of the optimization [4].

For an optimized problem

$$\min f(X)$$
$$s.t. \begin{cases} g_i(X) \geq 0, \ i = 1, 2, \ldots\ldots, l, \\ h_j(X) = 0, \ j = 1, 2, \ldots\ldots m, \end{cases} \tag{1}$$

When considered the objective function f (X) as an energy surface on the feasible set (solution space) which is uneven. If a smooth ball rolling on the surface, it may well stop because it dropped into a concave which may not the deepest. The simulated annealing method is similar to drive the ball with a horizontal force along the horizontal direction. If the force that effected on the ball is great enough and its concave is not really deep, the ball will roll out from this hollow to another concave by the horizontal force. If the horizontal force is appropriately controlled, the ball is most likely to stay in the deepest concave which is the global optimal solution or the approximate global optimal solution of the optimization. The horizontal force effected on the ball corresponding to the temperature (T) in the simulated annealing algorithm, horizontal force reduced corresponding to the temperature decreased [5, 6].

3 The Fundamental Iterative Procedure of the Simulated Annealing Algorithm

1. Set initial temperature T_0 and initial point X, calculate the function value f (X) of the initial point.
2. Generate a random disturbanceΔX and get a new point $X' = X + \Delta X$, calculate the function value $f(X')$ and difference value of the functions $\Delta f = f(X') - f(X)$.
3. If $\Delta f \leq 0$, accept the new point as the initial point in next simulation.
4. If $\Delta f > 0$, calculate the acceptable rate $p(\Delta f) = \exp(-\frac{\Delta f}{K-T})$, obeying the uniform distributive pseudo-random number in the interval [0,1]. If $P(\Delta f) \geq r$, accept the new point as the initial point as the initial point in next simulation, or give up the new point and still take the initial point [7].

4 Establish Mathematical Model

4.1 Optimization Variables and Constraint Conditions

The variables in this paper is the warship overall elements which should be optimized. It includes length of waterline (L_ω), width of waterline (B_ω), draft (T), depth (D), displacement (Δ), the number of (N_{ad}), (N_s), (N_e), (N_a). The constraints of variables are considered before building the objective function. Some fluctuation atmosphere is given according to the principle dimensions and the number of carrier-based aircraft of the parent ship before the optimizing, as formula (2). A statistical range is drawn by formula (3) based on the statistical results of ship form factors on world's aircraft carriers [8].

$$s.t. \begin{cases} L_\omega \subset 291 : 311(\text{m}) \\ B_\omega \subset 38 : 40(\text{m}) \\ D \subset 28.7 : 30.7(\text{m}) \\ T \subset 9.8 : 11.8(\text{m}) \\ \Delta \approx 79123 : 83123(\text{m}) \\ N_{ad} \subset 4 \sim 6 \\ N_s \subset 6 \sim 8 \\ N_e \subset 6 \sim 8 \\ N_a \subset 50 \sim 60 \end{cases} \quad (2)$$

$$s.t. \begin{cases} L/B \approx 7.10 : 8.10 \\ L/\nabla^{1/3} \approx 7.00 : 7.90 \\ B/T \approx 3.40 : 4.10 \\ C_b \approx 0.57 : 0.62 \\ \overline{GM} \approx 2.7 : 3.4 \end{cases} \quad (3)$$

4.2 Objective Function

A comprehensive consideration of the speed and powering performance, the sea keeping performance, the stability, the operational effectiveness and cost is taken into account. The total objective function is built up, which can be expressed as following (formula 4).

$$
\begin{cases}
\min f_1(\mathbf{X}) = \dfrac{P_0 \Delta^{2/3} U^3}{\Delta_0^{2/3} U_0^3} \\[2mm]
\max f_2(\mathbf{X}) = T_\Phi = 0.58 * \sqrt{\dfrac{B_\omega^2 + 4Z_g^2}{GM}} \\[2mm]
\max f_3(\mathbf{X}) = T_z = 2.4\sqrt{T} \\[2mm]
\min f_4(\mathbf{X}) = \left| \overline{GM} - \overline{GM}_0 \right| \\[2mm]
\max f_5(\mathbf{X}) = E = \omega_{E1} * N_{ad} + \omega_{E2} * N_s + \omega_{E3} * N_e + \omega_{E4} * N_a \\[2mm]
\min f_6(\mathbf{X}) = C = \omega_{C1} * N_{ad} + \omega_{C2} * N_s + \omega_{C3} * N_e + \omega_{C4} * N_a + \omega_{C5} * \Delta
\end{cases} \tag{4}
$$

The weight coefficient ω_i should be counted in the process of building the minimum objective function [9], which can be used to control the optimized result of overall elements. Then the weight coefficient is counted as well and the minimum objective function is quickly generated by overall element factor in the stage of preliminary design. It can be expressed as following (formula 5).

$$
\min Z(x) = \omega_1 * \frac{P - P_{\min}}{P_{\max} - P_{\min}} + \omega_2 * \frac{T_{\phi\max} - T_\phi}{T_{\phi\max} - T_{\phi\min}} + \omega_3 * \frac{T_{z\max} - T_z}{T_{z\max} - T_{z\min}}
$$

$$
+ \omega_4 * \frac{\left| \overline{GM} - \overline{GM}_0 \right|}{\left| \overline{GM}_{\max} - \overline{GM}_{\min} \right|} + \omega_5 * \frac{E_{\max} - E}{E_{\max} - E_{\min}} + \omega_6 * \frac{C - C_{\min}}{C_{\max} - C_{\min}} \tag{5}
$$

Therein, Δ is for displacement (t); U is for ship speed (kn); P is for effective power (hp); subscript 0 is for the index of parent ship Kitty Hawk [10]: set Δ as 81,123(t), U_0 as 32(kn), P_0 as 280,000(kn). With the expectation of the invariable main energy power, the speed can reach 33(kn) from 32 (kn). N_{ad} is the number of early warning, N_s is the number of antisubmarine warfare aircraft, N_e is the number of electronic warfare aircraft, N_a is the number of combat aircraft, ω_{Ei} is the weight of shipboard aircraft occupied in operational effectiveness, and set $\omega_{Ei} = 0.25$. ω_{ci} is the number of shipboard aircrafts and the amount of displacement, and set $\omega_{ci} = 0.20$.

Selection of weight value should be based on the importance of target. Considering the speed and powering performance, the sea keeping performance, the stability, the operational effectiveness and economy to be important equally, variables are set as follows $\omega_1 = 0.2$, $\omega_2 = 0.2$, $\omega_3 = 0.1$, $\omega_4 = 0.1$, $\omega_5 = 0.2$, $\omega_6 = 0.2$. Combined variable is generated in random within the fluctuation range based on the objective function of mathematical model as mentioned above, and

the results of variable (L_ω, B_ω, T, D, Δ, N_{ad}, N_s, N_e, N_a) are calculated in FORTRAN programming language.

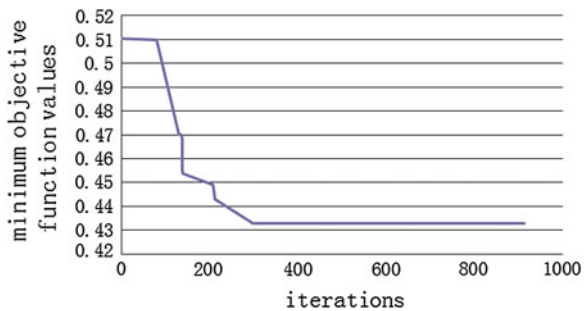## 5 The Optimized Result of Simulated Annealing Algorithm

The parameter selection of solving process is: t_{max} (initial temperature) $= 1 \times 10^6$. The times of inner cycle iteration at the same temperature is 1,000 times. The way of temperature reduction is to decrease the common ratio that is $T_{k+1} = 0.98T_k$; The termination criterion of major cycle is $T_k < 1 \times 10^{-4}$. Finally, the returned value is the optimized value $L\omega$, $B\omega$, T, D, Δ, N_{ad}, N_s, N_e, N_a. The temperature decreased 913 times and reached the termination criterion, which is known from the parameter selection, so the computer selected the optimal solution after $913 \times 500 = 456{,}500$ times iteration [11].

The procedure of aircraft carrier overall element is convergent and optimized results computed by simulated annealing algorithm are shown in the Table 1.

Table 1 Main factor and optimized results of minimum objective function

Iterations	Z_{min}	L_ω (m)	B_ω (m)	D (m)	T (m)	Δ (t)	N_{ad}	N_s	N_e	N_a
1	0.5105	297.3	39.3	29.6	10.5	80,289	5	7	7	59
80	0.5097	304.8	39.5	29.9	11.2	79,994	5	7	7	58
130	0.4704	302.5	39.3	30.1	11.1	79,990	5	7	7	59
139	0.4538	301.9	39.4	30.1	11.1	81,018	5	7	7	59
190	0.4504	297.8	39.4	30.2	10.9	79,909	5	7	7	59
207	0.4491	297.6	39.2	30.1	11.1	79,847	5	7	7	58
208	0.4477	297.9	39.4	30.2	11.1	79,945	5	7	7	59
212	0.443	298.3	39.3	30.3	11.1	80,055	5	7	7	58
298	0.4329	298.6	39.4	30.3	11.2	80,115	5	7	7	59
...	...	...	...	...	...	...	...	...	...	...
913	0.4329	298.6	39.4	30.3	11.2	80,115	5	7	7	59

Fig. 1 The optimized results of minimum objective function

Table 2 The comparison between the main factors optimized results and parent type

Indeterminate	Hawk	Simulated annealing
L_ω (m)	301	298.6
B_ω (m)	39	39.4
T (m)	10.8	11.2
D (m)	29.7	30.3
Δ (t)	81,123	80,115
N_{ad}	4	5
N_s	6	7
N_e	6	7
N_a	56	59

Table 3 The comparison between the main factors optimized results and parent type

	Hawk	Simulated annealing
P (hp)	28	27.76
T_Φ (s)	22.2	22.89
T_Z (s)	7.89	8.04
$\overline{GM}$ (m)	3.08	2.99

When the annealing process iterates to the 298 steps, the minimum function begins to steadily converge to $Z_{min} = 0.4329$ until the end of the iteration, which can be known from the optimized results. That is to say, the calculation system has obtained the best scheme (Fig. 1).

6 Comparative Analysis of Optimized Results and the Parent Ship

The comparison between the results of simulated annealing optimized method and the parent type is listed in Table 2.

Through the simulated annealing optimized method, the ultimate principal dimension factor value and four types of the number of shipboard aircrafts are concluded. According to the objective function value for the corresponding initial metacentric height $\overline{GM}$, effective power P and natural rolling period T_Φ and heave natural period T_z, and with the comparison of parent ship, the results are shown in Table 3.

Through the comparison with parent ship, the simulated annealing optimization method of the scheme have been improved in the speed and powering, initial stability, sea keeping, operational effectiveness and economy, which can be used as a quick generated optimization method in preliminary demonstration and design stage.

7 Conclusion

The results of optimization example indicate that the simulated annealing algorithm applied to the overall elements optimization of the aircraft carrier is feasible and effective. Its every iteration reflects the equilibrium of two strategies, concentration and diffusion. A variable is selected according to the last variable regularly, the objective function ultimately converge to $Z = 0.269$ until the end of iterative process. It is not influenced by the initial solution. There is also no requirement for the variable nature and convergence to the global optimized solution.

References

1. Huang P (2009) Optimal theories and methods. Tsinghua University Press, Beijing
2. Shao KW, Ma YY (2005) The ship technology and design introduction. National Defense Industry Press, Beijing
3. Sheng ZB, Liu YZ (2003) Theory of ship. Press of Shanghai Jiaotong University, Shanghai
4. Kirkpatrick S, Gelatl CD, Vecchi MP (1983) Optimization by simulated annealing. Science 220:671–680
5. Kang LS, Xie Y, You SY (1994) Non-numerical parallel algorithm- simulated annealing algorithm. Science Press, Beijing
6. Liu YD (2010) Principles of ship designing. National Defense Industry Press, Beijing
7. Ali MM, Storey C (1997) Aspiration based simulated annealing algorithm. J Glob Optim 11(2):181–191
8. Sun SN (2000) Modern aircraft carrier. Shanghai Popular Science Press, Shanghai
9. Ge L, Hou YL, Meng XY (2011) Design the principal dimensions of ships based on a fuzzy hybrid operator. ICIA. doi:10.1109/ICINFA.2011.5949115
10. Hai T (2006) US aircraft carriers CV-63 "The Kitty Hawk" [J]. Shipborne Weapons 11:32–37
11. Hou YH, Hu YL (2011) Group decision making of naval ship principal dimensions based on PSO algorithm. Adv Mater Res 201–203:1233–1237

Vehicle Scheduling Problem on Trees

Jie Zhou, Tianming Bu, Hong Zhu and Yixiang Chen

Abstract In this paper, we study a subproblem of vehicle scheduling problem, which is in general strongly NP-hard. In this problem, the road map is a tree, the release times and the handling times of all the tasks are zero, the deadlines of all the tasks are all the same, and the vehicle has to return to the root finally. We aim to find an optimal schedule such that the total value of the completed tasks are maximized. We show that this problem is NP-hard, and give a pseudo polynomial time algorithm for this problem.

1 Introduction

Vehicle scheduling problem (VSP, for short) has been studied as one of the most important scheduling problems [1–4]. Given a simple connected graph $G = (V, E)$ with vertex set $V = \{1, 2, \cdots, n\}$ and edge set E. For each edge $(i, j) \in E$, there is a

J. Zhou (✉) · T. Bu · H. Zhu · Y. Chen
Shanghai Key Laboratory of Trustworthy Computing, East China Normal University,
Shanghai 200062, China
e-mail: jiezhou@shnu.edu.cn

T. Bu
e-mail: jiezhou@shnu.edu.cn

H. Zhu
e-mail: jiezhou@shnu.edu.cn

Y. Chen
e-mail: jiezhou@shnu.edu.cn

J. Zhou
School of Mathematics and Physics, Shanghai Normal University, Shanghai 200234, China

H. Zhu
School of Computer Science, Fudan University, Shanghai, China

weight $w_{i,j}$, indicating the time needed for the vehicle to travel from vertex i to j. At each vertex i, there is a task (such as items to be picked up or facilities to be fixed), also written as i, which has four parameters: the release time r_i, the deadline d_i, the handling time h_i, and the value v_i. We have a vehicle to deal with the tasks. The tasks are nonpreemptive, that is, they must be processed without interruption. When a task is executed after its release time and completed before its deadline, the vehicle gets the value of the task. The aim of VSP is to work out a scheduling of the tasks and a routing for the vehicle to minimize (or maximize) a given objective function, such as the maximum tour time (i.e., the time when the vehicle returns to the initial node after it completes the last task in the schedule) of the vehicle, the makespan (i.e., the maximum completion time of tasks), the maximum lateness computed from deadlines, the total value of the completed tasks, and so on.

For VSP on graphs with the objective to maximize the total value of the completed tasks, Nagamochi et al in [5] give an $O(\rho n^{3+r})$ time algorithm that computes a 2-approximate solution to VSP, where $\rho = \max_{j,j'}(d(j) - r(j))/h(j')$ and r is the maximum number of tasks that can be processed by the vehicle after processing a task j and before visiting the processed task j again by deadline $d(j)$. The algorithm works efficiently on sparse graphs with time complexity $O(\rho n^2)$.

There are also a lot of practical applications of efficient algorithms for VSP on trees. For example, consider an automated guided vehicle or a handling robot that handles objects in a building with simple structures of corridors. Each object is placed in a room, and a vehicle is originally at the first floor to handle the objects. If each floor corresponds to a subtree and each room a leaf vertex, the road map of the vehicle forms a tree:
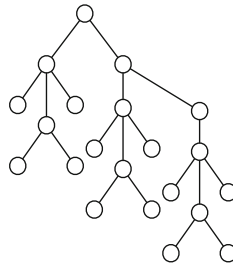
If the handling time is zero and the objective is to minimize the time when the vehicle returns to the root after completing all the tasks, Nagamochi et al. in [6] prove that this problem is strongly NP-hard by construct a reduction from 3-Partition which is known to be strongly NP-hard [7]. Karuno et al. in [6] give a 2-approximation for this problem with time $O(n \log n)$ which produces a scheduling satisfying the depth-first constraint.

Karuno et al. showed that when all jobs are available at time zero, the handling times are all zero, and the objective is to minimize the maximum lateness with respect to deadlines, the vehicle scheduling problem is strongly NP-hard [8].

We study VSP on trees with the objective to maximize the total values of the completed tasks in this paper. We show that if the release times and handling times

for all the tasks are zero, and all the tasks have a common deadline D, this problem is NP-hard, and give a pseudo polynomial time algorithm to this problem. Based on this pseudo polynomial time algorithm, we render a fully polynomial approximation scheme(FPTAS, for short) for this problem. When the release times of the tasks are not the same, or the handling times are not all zero, or the road map is a general graph, it can be proved that TREE-VSP-D becomes strongly NP-hard. When the deadline of the tasks are not the same, this problem is also strongly NP-hard.

2 Preliminaries

Strongly NP-hard and pseudo-polynomial time algorithm [7]. A problem is said to be strongly NP-complete, if it remains so even when all of its numerical parameters are bounded by a polynomial in the length of the input.

An algorithm runs in *pseudo-polynomial time* if its running time is polynomial in the size of the input and the numeric value of the input (which is exponential in the length of the input—its number of digits).

A problem can be proved to be *strongly NP-hard* by a pseudo-polynomial time reduction from a strongly NP-complete problem to it.

PTAS and FPTAS [9]. An algorithm A is said to be a *polynomial time approximation scheme (PTAS, for short)*, if it takes an instance of an optimization problem and a parameter $\varepsilon > 0$, and produces a solution that is within a factor $1 + \varepsilon$ of being optimal, with running time bounded by a polynomial in the size of instance I and ε, for each fixed $\varepsilon > 0$.

An algorithm A is said to be a *fully polynomial approximation scheme (FPTAS, for short)*, if it is a PTAS, and its running time bounded by a polynomial in the size of instance I and $\frac{1}{\varepsilon}$, for each fixed $\varepsilon > 0$.

3 Vehicle Scheduling Problem on Trees

We give the definition of the vehicle scheduling problem on trees (TREE-VSP, in short) as follows: Given a tree $T = (V, E)$, where $V = \{1, 2, \cdots, n\}$ and the root is denoted as 1. For each $(i, j) \in E$, there is a weight $w_{i,j}$, indicating the time needed for the vehicle to travel from vertex i to j. At each vertex i, there is a task, also written as i, which has four parameters: the release time r_i, the deadline d_i, the handling time h_i, and the value v_i. There is a vehicle, initially being at the root of the tree, to deal with the tasks. Task i is called *completed* if it is executed within the time interval $[r_i, d_i]$ and the vehicle gets the value v_i. After the vehicle completes the other tasks, it should return to the root by time D. To ensure this in the frame of the definition of VSP, let $h_1 = 0$, $r_1 = d_1 = D$ with $D \geq d_i$ for

$i = 2, \cdots, n$, and v_1 is large enough such that the root have to be completed to maximize the total value.

For each instance stated above, a *scheduling S* is defined as (J, σ), where $J \subseteq V$ is the set of completed tasks, $\sigma : \{1, 2, \cdots, |J|\} \to J$. For $i \in J$, $\sigma(i)$ is executed the *i*th. We seek nonpreemptive scheduling for TREE-VSP to maximize the total value of the completed tasks.

4 Complexity of the Problem

Theorem 1 If the release time and the handling time of each non-root task are all zero, and all the tasks share a common deadline *D*, *TREE-VSP is NP-hard.*

Proof We prove it by constructing a reduction from Knapsack Problem which is NP-hard. Given an instance *I* of Knapsack Problem as follows: A set $A = \{1, \cdots, n\}$ of *n* items and a bound *W*, each having a nonnegative weight w_i and a value v_i. The goal is to select a subset *S* of *A* with maximum total value $\sum_{i \in S} v_i$, subject to the restriction that $w_{i \in S} w_i \leq W$. Or equally, we can ask: for a given *v*, whether there exists a subset $S \subseteq A$ such that $\sum_{i \in S} v_i \geq v$, subject to $w_{i \in S} w_i \leq W$?

We construct the instance I' of TREE-VSP as follows: the road map is a tree $T' = (V', E')$ with $V' = A \cup \{0\}$, $E' = \{(0, i) | i \in V'\}$. For each $i \in A$, let $r'_i = h'_i = 0$, $v'_i = v_i$, $w'_{0i} = \frac{1}{2} w_i$ and $D = W$. 0 is the root of the tree and $h_0 = 0$, $r_0 = d_0 = D$, $v_0 = n$. It is not hard to show that there is a scheduling $S' = (J, \sigma)$ such that $val(S) \geq v + n$ for I' if and only if there is a solution *S* for *I*.

For the subproblem described in theorem 1, we denote it as TREE-VSP-D. It it is weakly NP-hard so that there are Pseudo Polynomial Algorithms for it. When the release times of the tasks are not the same, or the handling times are not all zero, or the road map is a general graph, it can be proved that TREE-VSP-D becomes strongly NP-hard.

5 A Pseudo Polynomial Algorithm for TREE-VSP-D

In this section, we give a pseudo polynomial algorithm for TREE-VSP-D by dynamic programming.

Definition 1 *I* is an instance of the problem TREE-VSP-D, and $S = (J, \sigma)$ is a scheduling for *I*. We say that *S* is a depth first scheduling, if *S* satisfies the following constraint: for each $i \in J$, once the vehicle reaches a vertex *i* from its parent in *T*, it cannot return to the parent unless it has completed all tasks both in *J* and the subtree rooted at *i*.

We write a a depth first scheduling as S^{DF}. By a depth first scheduling S^{DF}, each edge $(i,j) \in E$ is traversed at most twice. The following lemma states that the optimal scheduling for the instance of TREE-VSP-D satisfies the depth first property.

Lemma 1 Given an instance I of TREE-VSP-D, the optimal scheduling of I, $S^* = (J^*, \sigma^*)$, satisfies the following:

(1) If $i \in J^*$, then the father of i in T is also in J^*.
(2) S^* is depth first.

Proof Omitted.

Before proceeding to describe the algorithm further, we would like to fix some symbols here. Given a schedule S, we assume that the vehicle runs with no idle, and use $C(S)$ to denote the time when the vehicle finally returns to the root, $val(S) = \sum_{i \in S} v_i$ to denote the total value the vehicle gets.

Given an instance I of TREE-VSP-D, without loss of generality, assume that the vertexes (tasks) are given according to the pre-order traversal on the tree, that is, $V = \{1, 2, \cdots, n\}$, where node 1 is the root and i is visited ith by a depth first traversal. We construct a new instance I' as follows: Given the road map $T = (V, E)$ and the tasks V, for a given objective value $\mathbf{v}$, we want to compute a schedule S to minimize $C(S)$, satisfying that $val(S) \geq \mathbf{v}$. We denote this new problem as TREE-VSP-v.

If there is an algorithm for TREE-VSP-v, for $\mathbf{v} = 0, 1, \cdots, v^*$, where $v^* = \sum_{i=1}^{n} v_i$, we get a sequence of solutions $S_0^*, S_1^*, \cdots, S_{v^*}^*$. Then the optimal value of I can be obtained easily: it is v where v is the largest value such that $C(S_v^*) \leq D$.

Now we focus on the designing of a pseudo polynomial time algorithm for TREE-VSP-v using dynamic programming.

Let T^i represents the subtree of T, induced by vertex set $V' = \{1, \cdots, i\}$. We use f^i to be the father of node i in T^i (and hence also in T) and $f^+(i)$ the set of ancestors of i in T^i. Let $l = |f^+(i)|$, for $0 \leq k \leq l$, define

$$F_k(i) = \{f^k(i), f^{k+1}(i), \cdots, f^{l-1}(i), f^l(i)\}$$

Then we have that $F_0(i) = \{i, f(i), \cdots, 1\} = f^*(i)$, $F_1(i) = f^+(i)$, and $F_l(i) = \{1\}$.

Denote $T^i - i$ as the subtree of T^i by deleting the node i. Then $T^i - i = T^{i-1}$, that is, $i - 1$ is the largest among the numbers of leaves in $T^i - i$. Denote $F - \{i\} = F - i$.

The recurrence of $C(T^i, v, F)$ is given as follows:

- If $i \in F$, $C(T^i, v, F) = 2w_{i,f(i)} + C(T^{i-1}, v - v_i, F - i)$.
- Otherwise,
  $C(T^i, v, F) = \min\{2w_{f(i),i} + C(T^{i-1}, v - v_i, F \cup f^*(i),), C(T^{i-1}, v, F)\}$.

We can show that the recurrence is correct by which we can design an algorithm with time complexity $O(n^2 v^*)$ and it is pseudo polynomial for v^* is a number whose size might be exponential of n.

Based on the pseudo polynomial for TREE-VSP-v, we get an FPTAS for TREE-VSP-D as follows:

Algorithm 1 FPTAS for **TREE(V_{tour}, D)**$[0, D, 0, v]$

Given $\varepsilon > 0$, let $b = (\frac{\varepsilon}{n}) \max_i v(i)$.

For each $v(i)$, define $\widehat{v(i)} = \lceil \frac{v(i)}{b} \rceil$.

Using the dynamic algorithm described above to solve the TREE-VSP-D problem with values $\widehat{v(i)}$.

Return the schedule S.

Theorem 2 *Algorithm 1 runs in polynomial time for any fixed $\varepsilon > 0$.*

Theorem 3 *If S is the solution found by Algorithm 1, and S^* is any other solution satisfying $C^*_{max}(S^*) \leq D$, then we have $(1 + \varepsilon)Val(S) \geq Val(S^*)$.*

6 Conclusion

In this study, we investigate the complexity of a subproblem of vehicle scheduling which in general is strongly NP-hard and there is no efficient algorithm with good performance. When the release times and the handling times of all the tasks are zero, the deadlines of all the tasks are all the same, and the vehicle has to return to the root after completing other tasks, we show that this problem is NP-hard by a reduction from Knapsack Problem, and give a pseudo polynomial time algorithm for it.

When the deadlines of the tasks are not the same, this problem is strongly NP-hard. If we restrict the routing of the vehicle to be depth first, it is NP-hard. As a future work, We are going to make it out that if it is strongly NP-hard or not.

As for the instances of strongly NP-hard VSPs, we will focus on the approximation algorithms which also lies in the list of our future work.

References

1. Yu W, Liu Z (2011) Single-vehicle scheduling problems with release and service times on a line. Networks 57(2):128–134
2. Petersen HL, Larsen A, Madsen OBG (2012) Bjorn Petersen and Stefan Ropke, The Simultaneous Vehicle Scheduling and Passenger Service Problem, Transportation Science August 2012 trsc.1120.0429.
3. Bodin L, Golden B, Assad A, Ball M (1983) Routing and scheduling of vehicles and crews: The state of the art. Comput Oper Res 10:62–212
4. Bodin LD (1990) Twenty years of routing and scheduling. Oper Res 39:571–579
5. Nagamochi H, Ohnishi T (2008) Approximating a vehicle scheduling problem with time windows and handling times. Theor Comput Sci 393(1–3):133–146
6. Karuno Y, Nagamochi H, Ibaraki T (1997) Vehicle scheduling on a tree with release and handling times. Ann Oper Res 69:193
7. Garey M, Johnson D (1979) Computers and intractability: a guide to the theory of NP-completeness. W.H Freeman and Company, New York
8. Karuno Y, Nagamochi H, Ibaraki T (1996) Vehicle scheduling on a tree to minimize maximum lateness. J Oper Res Soc Jpn 39:345
9. Vazirani VV (2001) Approximation Algorithms. Springer-Verlag, Berlin

Reflections on the Enterprise to Change Management Concept: In Perspective of Concept Innovation

Nana Cao and Bin Lu

Abstract A successful enterprise must have a set of modern ideas. In the modern business system, the ideological concept is most important. With the rapid development of the world economy and science and technology, more and more people have increasingly recognized that the most needed of the outstanding enterprise is the continuous innovation of the ideas, not the profit index and not sophisticated computer systems. Based on the connotation of the concept and innovation, this paper analyzes more deeply the necessity to concept innovation of the enterprise, and puts forward several suggestions for an enterprise to change the management concept.

Keywords Concept · Innovation · Concept innovation

1 Introduction

With the advent of the 21st century, human society takes gradually from the industrial society into the knowledge society and into the new economic era which takes the innovation as one of the basic characteristics. So the enterprises and the individuals have to face the development of science and technology to change every day and be different every month and knowledge fission. The basis of the enterprise's survival and development is innovation, and the core is management innovation. Only we continuously promote the management innovation, the enterprise can remain invincible in domestic and international market competition. Therefore, the enterprises should continue to update their concepts, establish innovative consciousness, and gradually build their innovation system. Based on these measures, the enterprises can enhance their ability of management innovation and further improve their market competitive ability.

N. Cao (✉) · B. Lu
School of Economics and Management, Anhui University of Science and Technology,
Huainan city, People's Republic of China
e-mail: Lb815nn@163.com

2 The Connotation of the Concept Innovation

2.1 The Meaning of the Concept

The concept, that is recognition or thought, is that people formed the perception to objective things, and is a rational understanding of the objective world. It is playing a driving and guiding role to the behavior of people. Because there are numerous elements, the concept is an extremely complex system. At the same time, the concept has also a corresponding historical category. It is the product of under a certain era and condition, and often answers some problems that human need to be solved urgently. Thus it is easy to arouse people's psychological resonance, become widely accepted and recognized thought forms, and so form a relatively stable frame and structure of thinking. However, due to the universality of the concept, it has certain inertia, namely maintaining the status quo by the use of the existing concept.

2.2 The Meaning of Innovation

The term "innovation" was firstly proposed by American economist Schumpeter. The innovation, which Schumpeter referred to, is essentially the concept on economic significance. It means to the re-combination with the factors of production and production conditions [1]. Schumpeter believed that this concept included the following five situations: The introduction of a new product: that is, the new product or a new characteristic of a product which consumers are not familiar with; A new production method to be used: this method has not been verified by the experience in the manufacturing sector. It can exist in a new way to deal with a product commercially, not be established on the basis of new discoveries in science; A new market to be opened up: the manufacturing sector of a country has never entered into this new market, regardless of whether the market ever existed; A new supply source of raw materials or semi-manufactures to be plundered or controlled: wherever the source is already in existence or created at first; New organization of any kind of industry to be realized: such as creating a monopoly position or breaking a monopoly position.

2.3 The Connotation of the Concept of Innovation

Innovation is a multi-dimensional integration system, which includes five parts: technological innovation, institutional innovation, market innovation, concept innovation and management innovation. Among them, the concept and the concept innovation is the premise and pilot of innovation. No concept of innovation,

no activities of innovation. And the so-called concept innovation is the development of the concept and ideological revolution, namely all available knowledge and wisdom is applied to the existing products and services in order to create value. Thought innovation on creating market, which is the source of all innovation and the soul of enterprise development, requires that the enterprise keeps pace with the times, makes pioneering efforts, breaks the original thinking set and applies new thinking, new concept, new methods to operate businesses in order to create new value and meet new demands [2]. Therefore, the innovation activities are the concept innovation at first or carried out under the guidance of the concept innovation.

Certain concept forms certain management system of the enterprise and management mode of the enterprise, and certain management system of the enterprise and management mode of the enterprise causes some concept to form. The practice of enterprise management is always being under the guidance of certain concept. Due to a concept, the enterprise may be successful and obtain great development. However, the obstacles to further development of enterprises are precisely these concepts that have brought success to the enterprise. "Innovation is a destructively creative force" [3]. This damage will break the original balance of the enterprise, and the enterprise will pay more hardships and accompany with many risks for finding a new "cheese". But based on the diversity and variability of consumer preferences, the small and medium-sized enterprises must constantly provide marketable products according to the change of consumer demands. Only changes can meet changes with constancy and the enterprises get to survive and develop in changes. So if an enterprise wants to make great progress, it must constantly update their concepts and so as to gradually rise to the concept innovation in order to drive the enterprise to form the management system and the management innovation of the enterprise management mode [4].

3 The Necessity for Enterprise Concept Innovation

Building country should have their own foundation, and a man should have his own interpersonal skills. Certainly, run enterprises can not be short of one criterion, which is managing consciousness. The innovation of enterprise managing consciousness is the premise of other innovation. It is fundamental and overall to the impact on the development of enterprises. Enterprises adopting the different management methods, means and forms in different times and different contexts are the reaction to update consciousness. The concept innovation (management concept) oriented mode has great guiding significance for the enterprises to establish a correct social responsibility and good service consciousness. This mode is helpful for enterprises to choose appropriate management strategies and tactics.

3.1 Concept Innovation is the Life of Enterprise

Enterprise can not stick to conventions in modern society. Zhang Ruimin, Haier CEO, has said in his book—Haier is the sea: every step of Haier development accompanies with breakthroughs of the innovation and constantly going beyond itself, so innovation has become the soul of Haier culture. To the enterprise, no "incredible", no "impossible". As long as it has constantly innovation in the concept, management, systems and technology, "incredible" and "impossible" of things which people often said can be turned into reality and become the profits of the enterprise. These start with breakthrough "incredible" and "impossible", such as, Edison invented the light bulb, Gates engaged in Microsoft. In the fierce international competition, pressure and the sense of crisis has forced many enterprises to find the source of growth every day from a large number of "incredible" [5].

3.2 The Enterprise Concept Innovation is the Fundamental Guarantee for other Innovative Activities

The fundamental guarantee of the enterprises concept innovation for other innovative activities mainly reflected in the organizational system. In China, many scholars take technology innovation as the most important part of enterprise innovation and domestic research results on enterprise innovation is also reflected in the issue of technological innovation. Since the innovation theory emerged, there are two independent direction of development: technological innovation theory and institutional innovation theory. Among them, technological innovation theory has been paid more attention [6]. But in fact the enterprise innovation activities require the managers of the organization are beyond the concept and the theory firstly, and supplemented by the organizational structure and the innovation of the system in order to ensure that the entire organization adopts new technologies and new ways, which make innovation to become possible. Finally the enterprises provide new products and services for the society through management function activities, for examples decision-making, planning, command, organization, coordination and control [7]. Only enterprise, technological innovation is the microscopic foundation of the enterprise innovation system. It can be seen as the most direct innovatively behavior. And the concept innovation as the premise of enterprise organization system is the institutional guarantee of enterprise technology innovation.

3.3 The Concept Innovation has a Decisive Influence on the Enterprise Culture

The enterprise culture may lead to the innovative behaviors of the employees, develop their wisdom and potential, and activate each node in the business activities. It also plays the role on the regulation and the norms for the enterprise behavior. It is the internal condition of occurring corporate behavior. And behavioral characteristics are an important part of the enterprise culture [8]. So the concept innovation's influence on enterprise behavior is comprehensive and profound. The building of enterprise culture need realize the fundamental change of the enterprise managers' concepts and be really aware of the influence, which the enterprise management philosophy, the enterprise values, the enterprise spirit and the enterprise image is very important to the innovation and development. The enterprise can breed the sense of identity and the sense of belonging of the staff, and establish the stable and interdependent relationship between the staff and the enterprise. At the same time, the enterprise should achieve the self-management and the independent management of the employees in the corporate goal of reunification in order to realize the functions of the guidance, the condensation, the constraint and the incentive.

3.4 There are Still Some Problems to be Solved in the Enterprise Concepts

The traditional management concepts in China's most companies are more deeply rooted, mainly as follows: at first, the management target of the enterprise is not clear. The enterprise can not accurately make a choice on the immediate interests and the long-term interests, the local interests and the overall interests, the economic interests and social interests. Secondly, the relative lack in the enterprises is the spirit of independent innovation. Many enterprises' thoughts and ideas are conservative and backward. The innovative consciousness, the market consciousness, as well as legal consciousness is relatively weak. The third, the core competence of the enterprise has not fully cultivated. The organizational structure is not reasonable and the organizational system is not perfect, especially the less investment on science and technology, as well as the weak awareness of intellectual property rights. In addition, the building of enterprise culture is still remaining in the awareness level. The enterprises have not taken enterprise culture as a major business into the concrete action. Therefore, the enterprise must transform their ideas.

In China's current conditions, the reform and the innovation on the concepts and ideological emancipation has a decisive role in the development of enterprises. The idea innovation and the system innovation may be more urgent and important than the technological innovation. Therefore, the enterprises should activate the

concepts and the ideas of the enterprises management level and make their enterprises management level to maximize to accept the radiation of new knowledge and new concepts [9]. The government should vigorously develop the intermediary service organizations and erected a bridge to connect the enterprises with the theoretical circles in order to make the channel smoother between basic research on the theory and application research on the enterprise practices and to further promote the innovation of the enterprise's own. Only in these ways, Chinese enterprises are hopeful to further enhance the competitiveness; China's economic take-off is possible, and China's tomorrow will be a better day.

4 The Thought of Transforming the Enterprise Management Concept

A successful enterprise must have a set of the modern ideological concepts. In modern operating system, the ideological concepts are at the most important position. With the world economic situation and the rapid development of science and technology, people gradually come to realize that the first requirement of the outstanding enterprise is not a profit index, not a sophisticated computer system, but the continuous innovation of the ideological concepts [10]. Thus, the modern enterprises should make their concepts to be transformed in the following aspects.

4.1 To Transform from "Enterprise = Machine" to "Enterprise = Family"

In traditional concepts, an enterprise is a machine to create the profit and the value. The staff is "the parts of this machine" or "the screws". While these parts are not working, adding a little oil is a good method. For the staff, this "oil" is the financial incentive. The communication between the leadership and the staff is small; the staff feels very boring and hates their works. In modern concepts, the enterprises are a large family, the staff is a member of this large family and everyone is the protagonist of this family. So the communication between the leadership and the staff is increasing, and the enterprises provide the opportunities of the social contact and the exchanges for its members. So that, the staff can think that their works are more fun and it is beneficial to play to their creativities.

4.2 To Transform from "Mall = Battlefield" to "Mall = Ecosystem"

In traditional concepts, the malls is a battlefield, based on the size of the enterprise in which the different environmental systems and the enterprise comprehensive strength in which the different environmental systems, every enterprise develops the attack strategy, position defense strategy, flanking defense strategy or guerrilla strategy. The goal of the enterprise to develop these strategies is to beat and defeat the rivals, and their own strength can be maintained and expanded. However in modern concepts, the awareness of the modern enterprise should be established on the basis of the "biological chain". It means that industry and commerce is composed by a variety of symbiotic relationship. Every aspect of the biological chain should not be killing each other. They should collaborate with each other to obtain the greater common interests in entire chain. For the enterprises, they should jointly develop the untapped market. At the same time, the ways and the means should be paid attention to in the activities of the development. It is to establish a symbiotic relationship between enterprises. The competitors should not be taken as the enemies, but as the partners.

4.3 To Transform from "Incentive = Punishment" to "Incentive = Desire"

In traditional concepts, the employees often take the conventional incentive as an elusive thing, even as a punishment. Modern creative incentives should make the enterprise employees to feel that the efforts to complete the goal are voluntary, so as to strengthen the staff's sense of mission. If the employees believe that the goal can be achieved, their tasks can be completed. Then the consistency of the personal interests and the interests of the organization will be realized.

4.4 To Transform from "Management = Control" to "Management = Service"

The conventional wisdom is that the management equals to "control". As the managers, they think that they should be adept in controlling the right in their own hands and dividing their assigned the sphere of influence. The managers can not to listen to different voices. So the exclusiveness is very strong. In the modern management concepts, the management is a service, and the managers should properly disperse their rights in order to prevent the power abuse of the personal. If the enterprise leaders want to become the innovative leader, it is necessary to carry out the creative works by organizing and participating in a variety of innovative

activities. So it is to establish the enterprise environment, which the innovation is advocated and encouraged, within the enterprise. Due to these measures, it is helpful to inspire the creative objections.

4.5 To Transform from "Change = Pain" to "Change = Development"

In the traditional concepts, the change is a very painful thing, and it means that the change will bring the instability to the enterprise, or even unemployment. Nevertheless, the modern concepts of the enterprise management think that if there is not the conflict between short-term goals of the enterprise and long-term goals of the enterprise, the enterprise is necessary to change and maintain the mobility and flexibility of the post. Meanwhile, the enterprises should make the decision quickly and decisively. Then, when the enterprises pursue the good wishes, they should do to prepare for the worst. In short, only a change, the enterprises will grow.

References

1. Schumpeter J (1990) The theory of economic development: an inquiry into profits, capital, credit, interest, and the business cycle. China Commerce and Trade Press, Beijing
2. Wei W (1998) The management on the road to the knowledge economy. Jianghai Acad J 6:21–23
3. Zhu Q (2003) Discussion on the management innovation of the enterprise. J Hunan Econ Manage Coll 10
4. Yu S (2003) The status quo of the small and medium enterprises innovation and the analysis of the countermeasures. J Liaoning Univ Technol 2
5. Caihong Z (2011) The trend of management innovation in current China's enterprises. China's Foreign Trade 4:266–267
6. Fu J (1998) Technical innovation. Tsinghua University Press, Beijing 6
7. Li B (2003) Innovative thinking and enterprise development. Modern 2:120–121
8. Chunfeng J (2008) Adhere to what innovation concept of the corporate culture is. Theory Front 6
9. Jianchun K (2011) Analysis of the idea innovation of the enterprise management in the harmonious society. Chin Bus Trade 24:66–67
10. Xinwu P (2011) National Innovation: the changes of system and idea. J Beijing Admin Coll 6:82–86

Relaxed Two-Stage Multisplitting Algorithm for Linear Complementarity Problem

Ban-xiang Duan and Dong-hai Zeng

Abstract In this paper, the authors first set up relaxed two-stage algorithm for solving the linear complementarity problem, which is based on the two-stage splitting algorithm, parallel computation and the multisplitting algorithm. This new algorithm provides a specific realization for the multisplitting algorithm and generalizes many existing matrix splitting algorithms for linear complementarity problems and linear systems. And then, they establish the global convergence theory of the algorithm when the system matrix of the linear complementarity problem is an H-matrix, M-matrix, strictly or irreducibly diagonally dominant.

Keywords Linear complementarity problem · Two-stage multi-splitting · Relaxed iterative · Convergence theory

1 Introduction

The linear complementarity problem, abbreviated $LCP(q, A)$, is to find a vectors $z \in R^n$ such that

$$z \geq 0, \quad q + Az \geq 0, \quad z^T(q + Az) = 0 \tag{1}$$

B. Duan (✉)
College of Computer Engineering and Technology, Guangdong Institute of Science and Technology, Zhuhai 519090 Guangdong, People's Republic of China
e-mail: duanbanx@hotmail.com

D. Zeng
Guangzhou College, Guangdong Institute of Science and Technology, Guangzhou, 510640 Guangdong, People's Republic of China
e-mail: 609223667@qq.com

or to show that no such vector z exists. where $A \in R^{n \times n}$ is a given real matrix and $q \in R^n$ is a given real vector. In scientific computing and engineering applications, many problems can lead to the $LCP(q, A)$. For example, the free boundary problems for journal bearing lubrication, the network equilibrium problems and the contact problems, etc. see [1] and the references therein.

Many iterative methods for solving systems of linear equations can be described by means of matrix splitting [2–6]. Recently, Bai and Wurentuya proposed two-stage multisplitting method for solving linear equations $Ax = b$, see [7, 8]. To accomplish this, they split the matrix A as the two stage multi-splitting, i.e. let

$$A = B - C, B = E_i - F_i (i = 1, 2, \ldots, \alpha)$$

In this paper, by reformulating the $LCP(q, A)$ as an implicit fixed-point equation based on a two-stage multisplitting of the system matrix A, we establish a class of relaxed two-stage algorithm for solving large sparse linear complementarity problem. It is a kind of parallel iterative algorithm. The results obtained from those splitting iterations are combined to define the two stage multisplitting iterates. Thus, the algorithm may be effectively implemented on multiprocessors.

This paper is organized as follows. In the next section, we present some necessary notation and preliminary results. In Sect. 3 we describe a relaxed two-stage algorithm for the linear complementarity problem. The convergence of the relaxed two-stage algorithm is proved in Sect. 4.

2 Preliminaries Results

In this section we review some know results needed in Sects. 3 and 4. To formulate them we begin with some basic notation used throughout the remaining part of this paper.

Definition 2.1 [4] A matrix $A = (a_{ij}) \in R^{n \times n}$ is called an M-matrix if $a_{ij} \leq 0$ for $i \neq j$ and $A^{-1} \geq 0$. The comparison matrix $<A> = (\alpha_{ij})$ of a matrix $A = (a_{ij})$ is defined by $\alpha_{ij} = \begin{cases} |a_{ij}|, i = j \\ -|a_{ij}|, i \neq j \end{cases}$. A matrix A is called an H-matrix if $<A>$ is an M-matrix.

Definition 2.2 [8] Let $A, M, N \in R^{n \times n}$. We define $A = B - C$ as a splitting of the matrix A if B is nonsingular. We say that the splitting is an H-splitting if $<A> - |C|$ is an M-matrix, i.e., $(<A> - |C|)^{-1} \geq 0$; an H-compatible splitting if $<A> = |B| - |C|$.

Definition 2.3 [8] Let $A, B, C, E_i, F_i \in R^{n \times n}$, and suppose that for some satisfying: (1) $A = B - C, B = E_i - F_i$ for $i = 1, 2, \ldots, \alpha$. (2) A, B, E_i is nonsingular for $i = 1, 2, \ldots, \alpha$.(3) D_i is a diagonal matrix with nonnegative entries for $i = 1, 2, \ldots, \alpha$,

and $\sum_{i=1}^{\alpha} D_i = I$ ($n \times n$ identity matrix). Then the collection of triples (B, C, E_i, F_i, D_i) is called a two-stage multisplitting of A.

Lemma 2.1 [7] *Let $A \in R^{n \times n}$ be an H-matrix with positive diagonal elements. Then the linear complementarity problem* (1) *has a unique solution.*

Lemma 2.2 [4] *Let $A, B, C \in R^{n \times n}$ and $A = B - C$ be a splitting. If the splitting is an H-splitting, then both A and B are H-matrices and $\rho(B^{-1}C) \leq \rho(<B>^{-1})|C| < 1$. Moreover, if the splitting is an H-compatible splitting and A is an H-matrix, then it is an H-splitting.*

Lemma 2.3 [5, 6] *Let $A \in R^{n \times n}$, Suppose that the matrix A satisfies one of the following conditions: (1) A is an M-matrix. (2) A is strictly or irreducibly diagonally dominant. (3) A is symmetric and positive definite. Then A is an H-matrix.*

3 The Relaxed Two-Stage Multisplitting Algorithm

In order to achieve higher computing efficiency by making use of the information contained in the matrix A, based on the two-stage matrix splitting iteration method we present a relaxed two-stage multisplitting iterative method for solving the linear complementarity problem.

Algorithm 3.1
(*Relaxed Two-stage Multisplitting Iterative Algorithm*)

Step 1: Let $z^0, y^{0,0} \in R^n$ be an arbitrarily given vector, set $r = 0$.

Step 2: For a given z^r, let $y^{r,0} = z^r$, for each i, let $y^{r,l+1}$ be an arbitrary iterative solution of the subproblems

$$\begin{cases} y^{r,l+1} \geq 0, \\ q - Cz^r + E_i y^{r,l+1} - F_i y^{r,l} \geq 0, \\ (y^{r,l+1})^T (q - Cz^r + E_i y^{r,l+1} - F_i y^{r,l}) = 0. \end{cases} \tag{2}$$

If $y^{r,l+1}$ satisfies the prescribed termination rule, then set $z_k^r := y^{r,l+1}$.

Step 3: Let

$$z^{r+1} = \omega \sum_{i=1}^{\alpha} D_i z_k^r + (1 - \omega) z^r. \tag{3}$$

Step 4: If z^{r+1} satisfies the prescribed termination rule, then stop. Otherwise, set $r := r + 1$ and return to Step 2.

4 The Convergence of the Relaxed Two-Stage Multisplitting Algorithm

In this section, we study the convergence properties of the sequence $\{z^r\}$ produced by the iterative scheme introduced in Algorithm 3.1.

Lemma 4.1 *Let* $A = B - C, B = E_i - F_i (i = 1, 2, \ldots, \alpha)$, *where* A, B *and* $E_i (i = 1, 2, \ldots \alpha)$ *are H-matrix with positive diagonal elements respective and* $A = B - C, B = E_i - F_i (i = 1, 2, \ldots, \alpha)$ *be an H-splitting. Let* z^* *be a unique solution of the linear complementarity problem* (1)*. Then, for an arbitrary vector* $q \in R^n$ *and any starting vector* $z^0 \in R^n, z^0 \geq 0$, *the uniquely defined sequence of iterates* $\{z_k^r\}$ *and* $\{z^r\}$ *satisfies*

$$< E_i > |y^{r,l+1} - z^*| \leq |C||z^r - z^*| + |F_i||y^{r,l} - z^*|. \tag{4}$$

Proof To establish the expression (4), we first remark that $y^{r,l+1}$ is uniquely defined because $E_i (i = 1, 2, \cdots \alpha)$ is an H-matrix with positive diagonal elements (see Lemma 2.1). We verify (4) component by component. Consider an arbitrary index j and assume that

$$|y^{r,l+1} - z^*|_j = (y^{r,l+1} - z^*)_j$$

Under this assumption, the inequality (4) holds clearly if $y_j^{r,l+1} = 0$, because the j-th component of the left-hand vector in (4) is then nonpositive and the right-hand component is always nonnegative.

Now, suppose that $y_j^{r,l+1} > 0$. Then, according to $y^{r,l+1}$ be an arbitrary solution of the following linear complementarity problem (2), we have

$$(q - Cz^r + E_i y^{r,l+1} - F_i y^{r,l})_j = 0.$$

On the other hand, we also have

$$(q - Cz^* + E_i y^* - F_i y^*)_j \geq 0.$$

Subtracting the last two expressions and rearranging terms, we deduce

$$(q - Cz^r + E_i y^{r,l+1} - F_i y^{r,l})_j - (q - Cz^* + E_i y^* - F_i y^*)_j \leq 0.$$

Therefore

$$(E_i(y^{r,l+1} - z^*))_j \leq (C(z^r - z^*) + F_i(y^{r,l} - z^*))_j$$

which implies

$$< E_i > |y^{r,l+1} - z^*| \leq |C||z^r - z^*| + |F_i||y^{r,1} - z^*|$$

(Because E_i is an H-matrix with positive diagonal elements and $|y^{r,l+1} - z^*|_j = (y^{r,l+1} - z^*)_j$)

In a similar fashion, we may establish the same inequality (4) if $|y^{r,l+1} - z^*|_j = (z^* - y^{r,l+1})_j$. Consequently, the inequality (4) must hold.

Theorem 4.2 *Let A, B and $E_i (i = 1, 2, \ldots \alpha)$ be an H-matrix with positive diagonal elements and $A = B - C, B = E_i - F_i (i = 1, 2, \ldots, \alpha)$ be a two-stage multisplitting of A. Let $A = B - C, B = E_i - F_i (i = 1, 2, \cdots, \alpha)$ be two H-splitting, $0 < \omega < \frac{2}{1+\rho}$, z^* be a unique solution of the linear complementarity problem (1). Then, for any starting vector $z^0 \in R^n, z^0 \geq 0$, the infinite iteration sequence $\{z^r\}$ generated by Algorithm 3.1 converges to the unique solution of the linear complementarity problem (1).*

Proof Since $A = B - C, B = E_i - F_i (i = 1, 2, \ldots, \alpha)$ be two H-splitting and A, B and $E_i (i = 1, 2, \ldots \alpha)$ be an H-matrix with positive diagonal elements. It follows from Lemma 2.2 that $\rho(B^{-1}C) \leq \rho(<B>^{-1}|C|) < 1$. Let z^* be the unique solution of the linear complementarity problem (1), where the uniqueness comes from the assumption that A is an H-matrix with positive diagonal elements. Then from the equality (3) we have

$$z^{r+1} = \omega \sum_{i=1}^{\alpha} D_i z_k^r + (1 - \omega) z^r.$$

Therefore

$$z^{r+1} - z^* = \omega \sum_{i=1}^{\alpha} D_i(z_k^r - z^*) + (1 - \omega)(z^r - z^*).$$

Then by Lemma 4.1 we have

$$<E_i> |y^{r,l+1} - z^*| \leq |C||z^r - z^*| + |F_i||y^{r,l} - z^*|.$$

So

$$<E_i> |z_k^r - z^*| \leq |C||z^r - z^*| + |F_i||z_k^{r-1} - z^*|.$$

It follows that

$$|z^{r+1} - z^*| = |\omega \sum_{i=1}^{\alpha} D_i(z_k^r - z^*) + (1 - \omega)(z^r - z^*)|$$

$$\leq \omega |\sum_{i=1}^{\alpha} D_i(z_k^r - z^*)| + (1 - \omega)|(z^r - z^*)|$$

$$\leq \omega \sum_{i=1}^{\alpha} D_i(<E_i>^{-1}|C||z^r - z^*| + <E_i>^{-1}|F_i||z^{r-1} - z^*|) + (1 - \omega)|z^r - z^*|$$

And when $r \to \infty$ we have

$$|z^{r+1} - z^*| \leq (\omega \sum_{i=1}^{\alpha} D_i(<E_i>^{-1}|C| + <E_i>^{-1}|F_i|) + (1-\omega)I)|z^r - z^*|$$

Let

$$H_\omega = [\omega \sum_{i=1}^{\alpha} D_i(<E_i>^{-1}|C| + <E_i>^{-1}|F_i|)] + (1-\omega)I.$$

Then we have

$$H_\omega = [\omega \sum_{i=1}^{\alpha} D_i(<E_i>^{-1}|C| + <E_i>^{-1}|F_i|)] + (1-\omega)I$$

$$= \omega \sum_{i=1}^{\alpha} D_i[<E_i>^{-1}|C| + <E_i>^{-1}|F_i|] + (1-\omega)I$$

$$= \omega \sum_{i=1}^{\alpha} D_i[<B+F_i>^{-1}|C| + <E_i>^{-1}|F_i|] + (1-\omega)I$$

$$\leq \omega \sum_{i=1}^{\alpha} D_i[(<B> + |F_i|)^{-1}|C| + <E_i>^{-1}|F_i|] + (1-\omega)I$$

$$\leq \omega \sum_{i=1}^{\alpha} D_i[<B>^{-1}|C| + |F_i|^{-1}|C| + <E_i>^{-1}|F_i|] + (1-\omega)I$$

Since $A = B - C, B = E_i - F_i (i = 1, 2, \ldots, \alpha)$ be H-splitting, we can get $(<B> - |F_i|)^{-1} \geq 0$. Then $<B>^{-1} \leq |F_i|(i = 1, 2, \ldots, \alpha)$ and $<B>^{-1}|C| \leq |F_i|^{-1}|C|$.

Hence

$$||H_\omega|| = ||[\omega \sum_{i=1}^{\alpha} D_i(<E_i>^{-1}|C| + <E_i>^{-1}|F_i|)] + (1-\omega)I||$$

$$\leq \omega||\sum_{i=1}^{\alpha} D_i[<B>^{-1}|C| + |F_i|^{-1}|C| + <E_i>^{-1}|F_i|]|| + |1-\omega|$$

$$\leq \omega||\sum_{i=1}^{\alpha} D_i[2<B>^{-1}|C| + <E_i>^{-1}|F_i|]|| + |1-\omega|$$

$$\leq \omega \sum_{i=1}^{\alpha} D_i[||2<B>^{-1}|C||| + ||<E_i>^{-1}|F_i|||] + |1-\omega|$$

$$\leq \omega \sum_{i=1}^{\alpha} D_i[\varepsilon + 2\varepsilon(\rho + \varepsilon)] + |1-\omega|$$

$$= \varepsilon\omega(1 + 2\rho + 2\varepsilon) + |1-\omega|$$

Because of $0 < \omega < \frac{2}{1+\rho}$, we also get $|1 - \omega| < 1$. Moreover, if $\varepsilon > 0$ is small enough, we have $2\omega\varepsilon + \omega\rho + |1 - \varepsilon| < 1$. Hence, $\rho(H_\omega) \le \|H_\omega\| < 1$, which completes the proof.

Corollary 4.3 *Let A, B and $E_i(i = 1, 2, \ldots \alpha)$ be an H-matrix with positive diagonal elements and $A = B - C, B = E_i - F_i(i = 1, 2, \ldots, \alpha)$ be a two-stage multisplitting of A. Let $A = B - C, B = E_i - F_i(i = 1, 2, \ldots, \alpha)$ be H-compatible splitting, $0 < \omega < \frac{2}{1+\rho}, z^*$ be a unique solution of the linear complementarity problem* (1). *Then, for any starting vector $z^0 \in R^n, z^0 \ge 0$, the infinite iteration sequence $\{z^r\}$ generated by Algorithm 3.1 converges to the unique solution of the linear complementarity problem* (1).

Proof As $A = B - C, B = E_i - F_i(i = 1, 2, \ldots, \alpha)$ be H-compatible splitting and A, B and $E_i(i = 1, 2, \ldots \alpha)$ be an H-matrix with positive diagonal elements, it follows from Lemma 2.2 that $A = B - C, B = E_i - F_i(i = 1, 2, \ldots, \alpha)$ be H-splitting. The proof is then reduced to that of Theorem 4.2 when the conditions involved in Theorem 4.2 hold.

Corollary 4.4 *Suppose that the matrix A satisfies one of the following conditions*

(1) *A is an M-matrix.*
(2) *A is strictly or irreducibly diagonally dominant.*
(3) *$<A>$ is symmetric and positive definite.*

Then, for any starting vector $z^0 \in R^n, z^0 \ge 0$, the infinite iteration sequence $\{z^r\}$ generated by Algorithm 3.1 converges to the unique solution of the linear complementarity problem (1), provided that the conditions of Theorem 4.2 hold.

Proof For (1) it is obvious that A is an H-matrix. In the case of (2) the matrix $<A>$ is strictly or irreducibly diagonally dominant. This implies that $<A>$ is an M-matrix from Lemma 4.1 and thus A is an H-matrix. The assumptions in (3) guarantee that $<A>$ is a symmetric and positive definite matrix with the sign pattern of an M-matrix. Therefore, $<A>$ is an M-matrix and A is again an H-matrix. The proof is then reduced to that of Theorem 4.2 when the conditions involved in Theorem 4.2 hold.

5 Conclusion

In this paper, we first present a relaxed two-stage multisplitting algorithm for solving the linear complementarity problem. This algorithm provides a specific realization for the two-satge algorithms and the multisplitting algorithms and suits the requirements of high-speed parallel multiprocessor systems. Then the convergence of the algorithm is analyzed for M- matrices, strictly or irreducibly diagonally dominant matrices and symmetric and positive definite matrices.

References

1. Cottle RW, Pang JS, Stone RE (1992) The linear complementarity problem. Academic Press, San Diedo
2. O'Leary DP, White RE (1985) Multisplittings of matrix and parallel solution of linear systems. SIAM J Algebraic Discrete Methods 6:630–640
3. Frommer A, Szyld DB (1992) H-splittings and two stage iterative methods. Numer Math 63:345–356
4. Varga RS (1962) Matrix iterative analysis. Prentice–Hall, Englewood Cliffs
5. Ortega JM, Rheinboldt WC (1970) Iterative solution of nonlinear equations in several variables. Academic, New York
6. Frommer A, Mayer G (1989) Convergence of relaxed parallel multisplitting methods. Linear Algebra Appl 119:141–152
7. Bai ZZ, Huang TZ (1994) Accelerated overrelaxation methods for solving linear complementarity problem. J UEST China 23:428–432
8. Wurentuya HS, Guo PF (2012) Relaxed tow-stage multisplitting methods for linear systems. J Inner Mongolia Univ Nationalities 27(2):148–150

The Analysis Methods About Business Process of E-Commerce Based on the Petri Net

Ouyang Hong, Xian-wen Fang and Min Yu

Abstract With the development and popularity of the Internet, electronic commerce has become an important mean of business in the developing current-social and economic. Presented methods on business process management mainly analyzed the correctness of the workflow model. However, these methods can't easily uncover the underlying risks existing in the model. In this paper, the analyzing method about the underlying risks of the business process is presented based on the Petri net, and we analyze the specific example by the proposed method. The theoretic analysis and the example effect show the method is effective.

Keywords E-commerce · Petri net · Business process · Modeling

1 Introduction

As a result of the growing adoption of Business Process Management (BPM) technology, different stakeholders need to understand and agree upon the process models that are used to configure BPM systems. While a substantial amount of literature is devoted to this topic, there is no overview of the various mechanisms that exist to deal with managing complexity in (large) process models. As a result,

O. Hong · X. Fang (✉)
School of Science, Anhui University of Science and Technology,
Huainan 232001, People's Republic of China
e-mail: fangxianwen@hotmail.com

O. Hong
e-mail: oyh8909@163.com

M. Yu
Department of the Information Science, Anhui University of Science and Technology,
Huainan 232007, People's Republic of China

it is hard to obtain an insight into the degree of support offered for complexity reducing mechanisms by state-of-the-art languages and tools [1].

Recently, Great attention has been paid to discuss the security of payment system in electronic commerce, analyze the basic structure and transaction process [2] of secure electronic transaction protocounderlying risks, and the transaction process of SET protocol is modified according to the actual situation of development in electronic commerce [3]. A solution of online secure payment system based on third party payment authority is presented [4]. Payment function is separated from merchant system. The supports in online payment is provided for merchants as independent third party payment authority. The framework is proved to be valid from both static and dynamic aspects, and can be used as a guider of system development [5]. Latest researches [6] discovered the application of RPN in workflow mining, and a mining algorithm based on Rough Petri Net [7] was proposed to discovery workflow process which integrates the features of rough set theory and Petri net. There also analyzed the reasonable verification theorems and algorithms of exist workflow network, proposes a proved rationality validation algorithm based on Petri net [8].

With the development of Internet technology, online shopping has become an important part of people's daily lives. E-commerce payment services mainly based on the third-party payment, however, because of the security vulnerabilities existed in the platan of the third-party payment system, which caused a loss to the third-party payment agencies and users. Petri net is not only capable of static description of the system, but also to reflect the behavior and properties of the system through the dynamic interaction, integrated data flow, control flow and state transition, which can easily describe the distribution of the system, resource sharing, synchronous, asynchronous, conflict, and other important characteristics.

2 Formal Preliminaries

This section clarifies notions and notations used in the remainders.

2.1 Petri Net

Petri net are graphic symbol-based modeling method which can be systematically analyzed by simulation, meanwhile there exists much mathematical theory foundation making Petri net with its extensions popular among researchers.

Definition 2.1   A marking Petri net can be defined as a four tuple $\Sigma = (S, T, F, M_0)$ where S is a finite set of places, T is a finite set of transitions, $S \cap T = \phi$ and $S \cup T \neq \phi$, $F \in (S \times T) \cup (T \times S)$ is a finite set of directed arcs, M_0 is an initial marking.

1. for $t \in T$, if $\forall s \in S : s \in {}^{\bullet}t \to M(s) \geq 1$, then t was enabled in the marking M, i.e. M[t>.
2. If M[t>, then t can fire in the marking M, M' is a new marking from M after t fired, i.e. M[t > M' for $\forall s \in S$, $M'(s) = \begin{cases} M(s)-1, s \in {}^{\bullet}t-t^{\bullet} \\ M(s)+1, \ s \in t^{\bullet}-{}^{\bullet}t \\ M(s), \quad else \end{cases}$

2.2 Methods to Analyze the Model of Petri Net

Definition 2.2 (*reachable marking graph*) Let $\Sigma = (S, T; F, M_0)$ be a bounded Petri net, the reachable marking graph of Σ is a 3 tuple $RG(\Sigma) = (R(M_0), \ E, \ P)$,

$$E = \{ \ (M_i, M_j) | M_i, M_j \in R(M_0), \exists t_k \in T : M_i[t_k > M_j) \}$$
$$P : E \to T, \ P(M_i, M_j) = t_k \Leftrightarrow M_i[t_k > M_j$$

$R(M_0)$ is the set of $RG(\Sigma)$'s vertex, E is the set of $RG(\Sigma)$'s arcs; if $P(M_i, M_j) = t_k, t_k$ is the side marking of arc $(M_i, \ M_j)$.

Definition 2.3 (*state equation*) Let $\Sigma = (S, T; F, M)$ be a marking Petri net, M_0 is an initial marking, A is the incidence matrix of Σ, if $M \in R(M_0)$, then there exist an n-dimensional matrix X, there establish state equation as below:

$$M = M_0 + A^T X$$

3 Modeling and Analysis of Business Process

3.1 Business Process of the Third-Party Payment

In the business process based on the third-party payment, firstly, the buyers get in touch with the seller to reach an agreement after selecting goods online, then the buyer remitted the payment to the third-party payment center, payment center as an intermediary party immediately notify the seller to deliver. After the buyer has received the goods then the payment center will remit the payment to the seller's account to complete the transaction. The mainly business process of e-commerce show in Fig. 1.

3.2 Modeling of Business Process

We would now like to establish model of the business process based on Petri net, showing in Fig. 2a. According to the established model, there is same problems

Fig. 1 Business process of
e-commerce
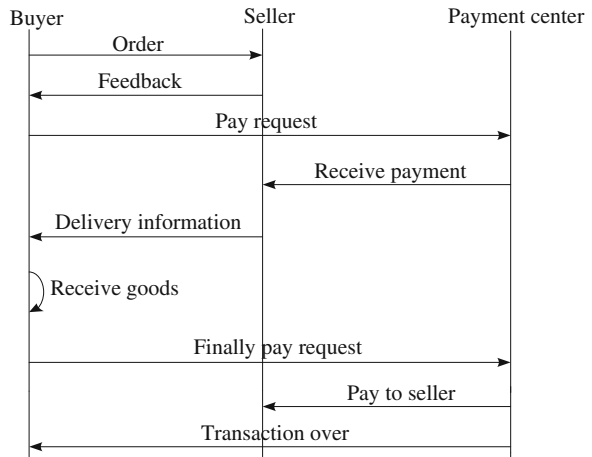
can be find. Until the transaction has finished, the sellers discovers the total was
wrong, but it is unable to prevent the loss. we add an arc to check the amount in
Fig. 2b,which can prevent the loss.

(a) Modeling on Petri net

(b) Modeling on Petri with checking
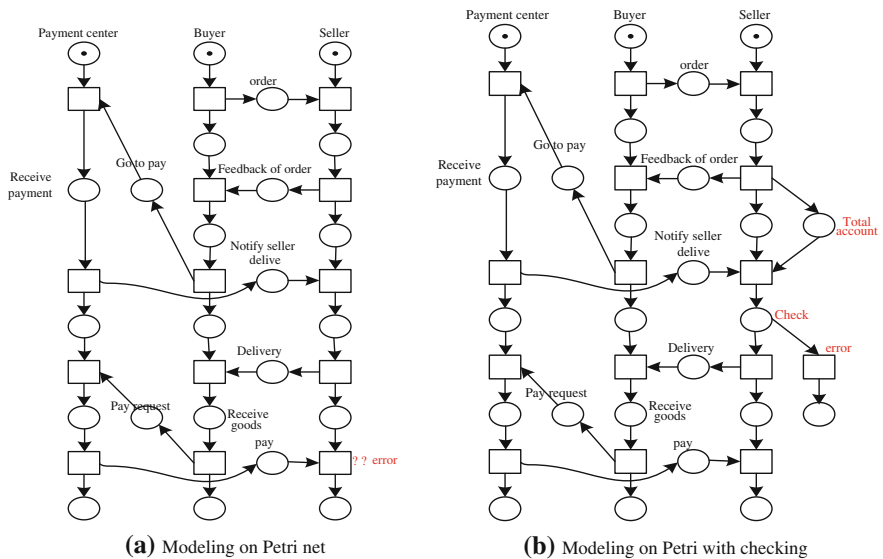
Fig. 2 Modeling on Petri net

3.3 Analysis of the Model

In Fig. 3, we label the transition and place of the model in Fig. 2b by T and S. According to the definition of reachable marking graph, we can drew the reachable marking graph of the labeled model, show in Fig. 3. With the reachable marking graph, we analyze characters of system, such as reachability, boundedness and liveness. The characters of network is corresponding to the nature of business process, the correctness of business process can be certificated by the simulation analysis of the dynamic characteristics of the network.

According to Fig. 3, the initial marking of the reachable marking graph of network is $M_0 = \{1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0\}$. Token transform from s_2 to s_4 and s_5 after t_1 fired. The sequence of marking change from M_0 to $M_1 = \{1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0\}$. Analogously, when t_2, t_3 fired, we can get markings such as M_2, M_3. When it comes to $M(s_1) = M(s_2) = \ldots = M(s_{23}) = 0, M(s_{24}) = M(s_{25}) = M(s_{26}) = 1$, in other word, when it comes to M_{15}, the process would finish without error. But if it comes to M_9 along with the transition t_9 firing, that means the system go into error (Figs. 4, and 5).
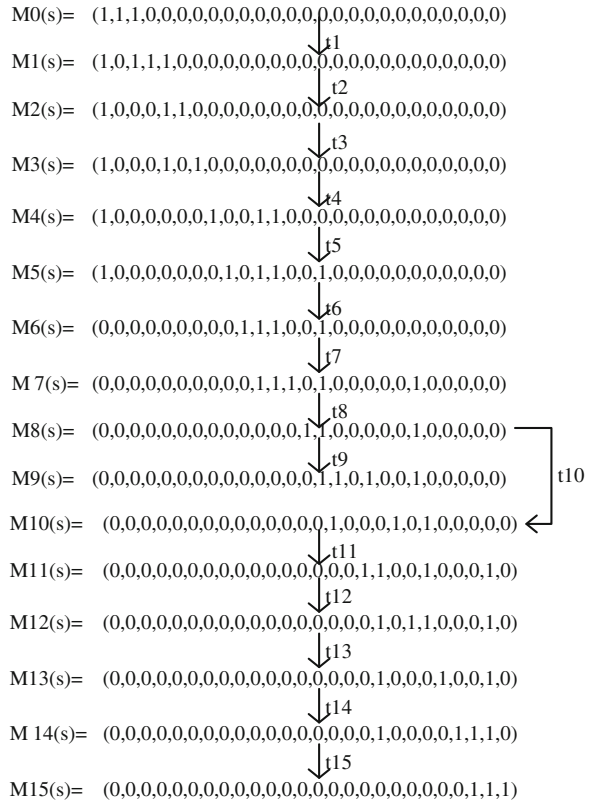

Fig. 3 Labeled model
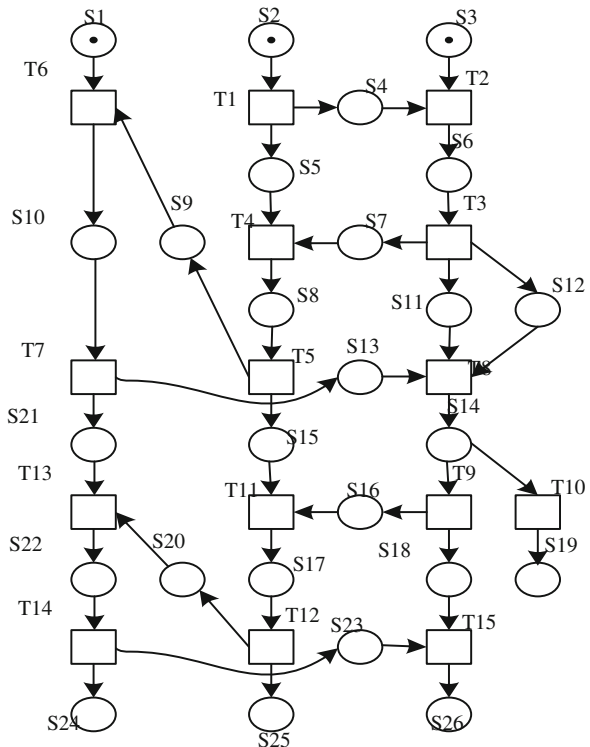
	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	s13	s14	s15	s16	s17	s18	s19	s20	s21	s22	s23	s24	s25	s26
t1		-1		1	1																					
t2			-1	-1		1																				
t3						-1	1					1	1													
t4					-1		-1	1																		
t5								-1	1								1									
t6	-1								-1	1																
t7										-1				1							1					
t8											-1	-1	-1		1											
t9															-1	1		1								
t10															-1				1							
t11														-1		-1	-1									
t12																		-1				1			1	
t13																						-1	-1	1		
t14																				-1				1	1	
t15																				-1				-1		1

Fig. 4 The incidence matrix of network

Fig. 5 The reachable marking graph of network

M0(s)= (1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0)

↓t1

M1(s)= (1,0,1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0)

↓t2

M2(s)= (1,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0)

↓t3

M3(s)= (1,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0)

↓t4

M4(s)= (1,0,0,0,0,0,0,1,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0)

↓t5

M5(s)= (1,0,0,0,0,0,0,0,1,0,1,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0)

↓t6

M6(s)= (0,0,0,0,0,0,0,0,0,1,1,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0)

↓t7

M 7(s)= (0,0,0,0,0,0,0,0,0,0,1,1,1,0,1,0,0,0,0,0,1,0,0,0,0,0)

↓t8

M8(s)= (0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,1,0,0,0,0) ┐

↓t9   │ t10

M9(s)= (0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0,1,0,0,1,0,0,0,0,0)   │

M10(s)= (0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,1,0,1,0,0,0,0,0) ◄─┘

↓t11

M11(s)= (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,1,0,0,0,1,0)

↓t12

M12(s)= (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,1,1,0,0,0,1,0)

↓t13

M13(s)= (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,1,0,0,1,0)

↓t14

M 14(s)= (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,1,1,1,0)

↓t15

M15(s)= (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1)

4 Conclusion

We analyzed the risks of the business process by establishing the model of Petri net on the business process which based on the third-party payment, and utilized verify method to analyze characters of the network. Compared with other modeling methods, Petri net is not only capable of static description of the system, but also to reflect the behavior and properties of the system through the dynamic interaction.

References

1. La Rosa M (2011) Managing process model complexity via abstract syntax modifications. Ind Info, IEEE Trans 7:614–629
2. van der Aalst Wil MP, Weijters T, Maruster L (2004) Workflow mining: discovering process models from event logs. IEEE Tran Knowl Data Eng 16:1128–1142
3. Sergey S, Matthias W, Jan M (2012) Business process model abstraction based on synthesis from well-structured behavioral profiles. Int J Coop Info Syst 21:55–83
4. Weixiang S, Tao L, Wei P, Tong S (2006) Incremental workflow mining with optional patterns systems. In: Weixiang S, Tao L, Wei P, Tong S 2006 IEEE conference on systems, man, and cybernetics. IEEE, Piscataway, USA, pp 2764–2771
5. Chunqin G, Huiyou C, Yang Y (2007) Overview of workflow mining technology granular computing. In: Chunqin G, Huiyou, C, Yang Y. GRC 2007. IEEE International Conference. IEEE, Piscataway, USA, pp 347–347
6. Vanhatalo J, Völzer H, Leymann F (2007) Faster and more focused control-flow analysis for business process models through sese decomposition. In: Vanhatalo J, Völzer H, Leymann F. Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics. Springer, Heidelberg, pp 43–55
7. Weidlich M, Elliger F, Weske M (2011) Generalised computation of behavioural profiles based on Petri-net unfoldings. Web Serv Formal Methods. doi:10.1007/978-3-642-19589-1_7
8. Vanhatalo J, Völzer H, Leymann F, Moser S (2008) Automatic workflow graph refactoring and completion. Service-oriented computing. 6th International conference. doi:10.1007/978-3-540-89652-4_11

Associated Relaxation Time for an Optical Bistable System with Coupling Between Non-Gaussian and Gaussian Noise Terms

Bing Wang and Xiuqing Wu

Abstract The relaxation time T of an optical bistable system with multiplicative non-Gaussian and additive Gaussian noise terms is investigated. The results indicate that the decay rate of fluctuation changes from slowing down to speeding up with increasing of non-Gaussian noise correlation time τ. But with increasing of the parameter q, the decay rate changes from speeding up to slowing down. The additive noise intensity D and correlation intensity λ slow down the fluctuation decay in the stationary state of the system.

Keywords Optical bistable system · Associated relaxation time · Non-Gaussian noise · Gaussian noise

1 Introduction

In recent years, the effects of noise on nonlinear systems has been widely investigated in many fields. Han et al. studied the laser intensity power spectrum and signal-to-noise ratio for a gain-noise model driven by colored cross-correlated pump noise [1]. Cao et al. invested the stochastic resonance of periodically driven system with multiplicative white noise and periodically modulated additive white noise [2]. In all these works, the fluctuations are assumed to be Gaussian noises. However, some recent experimental and theoretical results for one kind of crayfish and for rat skin offered strong indications, that there could be non-Gaussian noise sources in these nonlinear systems [3, 4]. Goswami et al. studied the barrier

B. Wang (✉) · X. Wu
Department of Physics and Mathematics, Anhui University of Science and Technology,
Huainan 232001, People's Republic of China
e-mail: hnitwb@163.com

X. Wu
e-mail: wxq328@163.com

crossing dynamics in presence of non-Gaussian noises and observed that non-Gaussian noise can induce resonant activation [5].

In the recent past, optical bistability has attracted a great amount of interest and has given rise to numerous experimental and theoretical studies. Du et al. investigated the normalized correlation function of an optical bistable system driven by cross-correlated white noises [6]. Wang et al. studied the effects of correlation between colored noises on the correlation function of an optical bistable system [7]. In all these works, the fluctuations assumed to be Gaussian noises. In this paper, the relaxation time of an optical bistable system with multiplicative non-Gaussian and additive Gaussian noise terms is studied.

2 The Steady-State Probability Density of the Optical Bistable System

The model for purely absorptive optical bistability in an optical cavity was introduced by Bonifacio and Lugiato [8]. The corresponding evolution is given as

$$\frac{dx}{dt} = y - x - \frac{2cx}{1 + x^2}, \tag{1}$$

c is the cooperativity parameter, y is the amplitude of the input light, x is the transmitted light amplitude. For large c, the input–output characteristics show the bistability. For large c and chosen $y = y_0$ within the regime of bistability, consider fluctuations of y and c. The corresponding Langevin equation reads

$$\frac{dx}{dt} = y_0 - x - \frac{2cx}{1 + x^2} + \frac{2x}{1 + x^2}\xi(t) + \eta(t), \tag{2}$$

$\xi(t)$ is a multiplicative non-Gaussian noise term

$$\dot{\xi} = -\frac{1}{\tau}\frac{d}{d\xi}V_q(\xi) + \frac{1}{\tau}\varepsilon(t), \tag{3}$$

with

$$V_q(\xi) = \frac{Q}{\tau(q-1)}\ln[1 + \frac{\tau}{Q}(q-1)\frac{\xi^2}{2}], \tag{4}$$

$\varepsilon(t)$ and $\eta(t)$ are Gaussian white noise, and the statistical properties are

$$\langle \varepsilon(t)\varepsilon(t')\rangle = 2Q\delta(t - t'), \tag{5}$$

$$\langle \eta(t)\eta(t')\rangle = 2D\delta(t - t'), \tag{6}$$

$$\langle \varepsilon(t)\eta(t')\rangle = 2\lambda\sqrt{DQ}\delta(t - t'), \tag{7}$$

Q and D are the intensities of $\varepsilon(t)$ and $\eta(t)$. τ is the correlation time of $\xi(t)$. λ is the correlation intensity between $\varepsilon(t)$ and $\eta(t)$. q in Eq. (4) denotes the departure from

the Gaussian noise. In the limit of $q \to 1$, the process $\xi(t)$ coincides with the Gaussian colored noise with self–correlation time τ

Applying the path integral approach, one has

$$\dot{\xi} = -\frac{1}{\tau_{eff}}\xi + \frac{1}{\tau_{eff}}\varepsilon_{eff}(t), \tag{8}$$

with

$$\langle \varepsilon_{eff}(t)\varepsilon_{eff}(t') \rangle = 2Q_{eff}\delta(t - t'), \tag{9}$$

$$\langle \eta(t)\varepsilon_{eff}(t') \rangle = \langle \eta(t')\varepsilon_{eff}(t) \rangle = 2\lambda\sqrt{Q_{eff}D}\delta(t - t'), \tag{10}$$

the effective noise correlation time $\tau_{eff} = 2(2 - q)\tau/(5 - 3q)$, and the associated effective noise intensity $Q_{eff} = (2(2 - q)/(5 - 3q))^2 Q$.

Using the Fox approach [9], the approximate Fokker–Planck equation corresponding to Eq. (2) can be written as

$$\frac{\partial P(x,t)}{\partial t} = L_{FP}P(x,t), \tag{11}$$

$$L_{FP} = -\frac{\partial}{\partial x}F(x) + \frac{\partial^2}{\partial x^2}G(x). \tag{12}$$

The drift coefficient $F(x)$ and the diffusion coefficient $G(x)$ are given by

$$F(x) = f(x) + 4\frac{Q_{eff}x(1 - x^2)}{C(\tau_{eff})(1 + x^2)^3} + 2\lambda\sqrt{Q_{eff}D}\frac{1 - x^2}{(1 + x^2)^2}, \tag{13}$$

$$G(x) = 4\frac{Q_{eff}x^2}{C(\tau_{eff})(1 + x^2)^2} + 2\lambda\sqrt{Q_{eff}D}\frac{x}{1 + x^2} + D, \tag{14}$$

$$f(x) = y_0 - x - \frac{2cx}{1 + x^2}, \tag{15}$$

the steady-state probability density of Eq. (11) is

$$P_{st}(x) = N\frac{1}{G(x)}\exp(\int_x \frac{F(z)}{G(z)}dz). \tag{16}$$

N is the normalization constant and if $x \le x_s, C(\tau_{eff}) = 1 - \tau_{eff}f'(x_{s1})$, if $x > x_s, C(\tau_{eff}) = 1 - \tau_{eff}f'(x_{s2})$. x_{s1}, x_s and x_{s2} are the steady-state values ($x_{s1} < x_s < x_{s2}$).

According the Eq. (16), the effects of q and τ on $P_{st}(x)$ can be studied with the numerical computation.

Figure 1 shows $P_{st}(x)$ as a function of x with different q. It is seen that there are two peaks in $P_{st}(x)$, the left peak is narrower than the right one. We can also find that the larger q is, the lower left peak is, and the higher right one is.

Fig. 1 $P_{st}(x)$ as a function of x with different values of $q, y_0 = 14.2, c = 20.02, D = 0.32, Q = 0.3, \tau = 0.5$
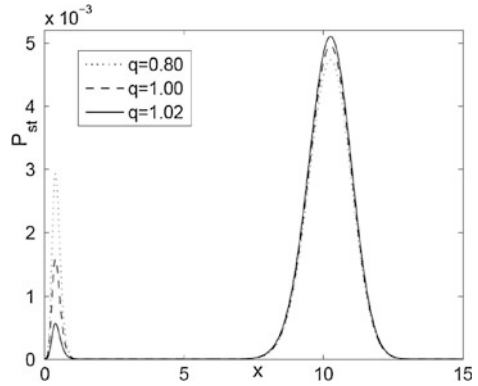
Fig. 2 $P_{st}(x)$ as a function of x with different values of $\tau, y_0 = 14.2, c = 20.02, D = 0.32, Q = 0.3, q = 0.8$
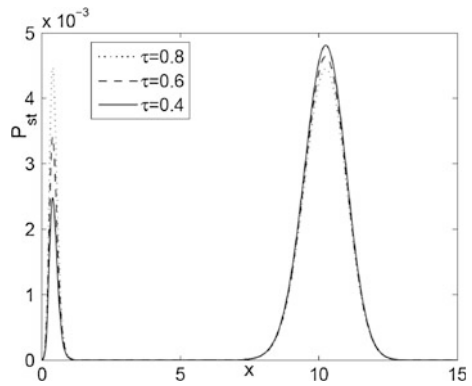

Figure 2 shows $P_{st}(x)$ as a function of x with different value of the correlation time τ. It is clear that the larger τ is, the higher left peak is, but the lower the right one is. From Figs. 1 and 2, we can find that q and τ play opposite roles on $P_{st}(x)$ of the system.

3 The Associated Relaxation Time of the Optical Bistable System

For a nonlinear stochastic system, the relaxation time T is defined as

$$T = \int_0^\infty C(s)ds, \tag{17}$$

$C(s)$ is the normalized correlation function and defined as

$$C(s) = \frac{\langle \delta x(t+s)\delta(t)\rangle_{st}}{\left\langle (x(t) - \langle x(t)\rangle)^2 \right\rangle_{st}}. \tag{18}$$

Fig. 3 T as a function of τ
with different values of
Q, $y_0 = 14.2$, $c = 20.02$,
$D = 0.32$, $\lambda = 0.5$, $q = 0.8$
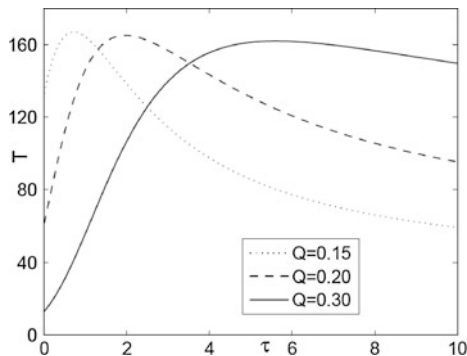
In terms of the adjoint operator L_{FP}^{+} in Eq. (12) and using the projection operator method, we get the approximation expression for the relaxation time

$$T = \left[\mu_0 + \frac{\eta_1}{\mu_1} \right]^{-1} \qquad (19)$$

$$\mu_0 = \frac{\langle G(x) \rangle_{st}}{\left\langle (\delta x)^2 \right\rangle_{st}}, \qquad (20)$$

Fig. 4 T as a function of q:
(a) $y_0 = 14.2$, $c = 20.02$,
$c = 20.02$, $\tau = 0.5$,
$D = 0.144$, $Q = 0.4$;
(b) $y_0 = 14.2$, $c = 20.02$,
$\tau = 0.5$, $\lambda = 0.4$, $Q = 0.4$
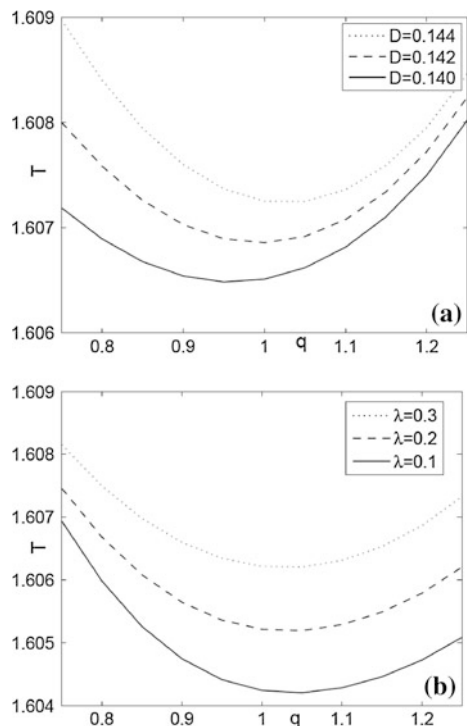
$$\eta_1 = \frac{\langle G(x)F'(x)\rangle_{st}}{\left\langle (\delta x)^2 \right\rangle_{st}} + \mu_0^2, \tag{21}$$

$$\mu_1 = -\frac{\left\langle G(x)[F'(x)]^2 \right\rangle_{st}}{\eta_1 \left\langle (\delta x)^2 \right\rangle_{st}} + \frac{\mu_0^3}{\eta_1} - 2\mu_0. \tag{22}$$

According to Eq. (19), the effects of the parameter q and the correlation time τ on relaxation time T can be analyzed by numerical computations.

Figure 3 displays the effects of τ on T. It is found that T has a maximum with increasing of τ. This means that with the increasing τ, the decay rate of the fluctuation turns over, from slowing down to speeding up. We also can find that with the increases of Q, the peak of the $T - \tau$ curve moves to right.

Figure 4 displays the effects of q on T. It is clear that q and τ play opposite roles on T. There exists a minimum in the $T - q$ curve, which indicates that with increasing q, the statistical fluctuation decay rate changes from speeding up to slowing down. We can also find T increases with increasing D or λ, in other words, D and λ slow down the fluctuation decay of the optical bistable system in the stationary state.

In summary, we have investigated the effects of the parameter q and non-Gaussian noise correlation time τ on the relaxation time T of an optical bistable system. We find that T has a maximum with increasing τ. However, T has a minimum with increasing of q. T increases with increasing additive noise intensity D or the correlation intensity λ.

References

1. Han LB, Cao L, Wu DJ, Wang J (2006) Stochastic resonance for bias signal-modulated noise in a single-mode laser. Phys A 366:159–166
2. Cao L, Wu DJ (2006) Stochastic resonance in a periodically driven linear system with multiplicative and periodically modulated additive white noises. Phys A 376:191–198
3. Borland L (1998) Ito-Langevin equations within generalized thermostatistics. Phys Lett A 245:67–72
4. Fuentes MA, Wio HS, Toral R (2002) Effective Markovian approximation for non-Gaussian noises: a path integral approach. Phys A 303:91–104
5. Goswami G, Majee P, Ghosh PK, Bag BC (2007) Colored multiplicative and additive non-Gaussian noise-driven dynamical system: mean first passage time. Phys A 374:549–558
6. Du LC, Mei DC (2008) Intensity correlation function of an optical bistable system subjected to cross-correlated noises. Phys Lett A 372:5529–5533
7. Wang B, Yan S (2009) Correlation function of an optical bistable system with cross-correlated additive white noise and multiplicative colored noise. Chin Opt Lett 7:838–840
8. Bonifacio R, Lugiato LA (1978) Optical bistability and cooperative effects in resonance fluorescence. Phys Rev A 18:1129–1144
9. Fox RF (1986) Rniform convergence to an effective Fokker-Planck equation for weakly colored noise. Phys Rev A 34:4525–4527

Bicyclic Graphs with Nullity $n-5$

Tian-tian Zheng

Abstract Let G be a simple undirected graph on n vertices, $A(G)$ be its adjacency matrix. The nullity $\eta(G)$ of the graph G is the multiplicity of the eigenvalue zero in its spectrum. In this paper, we characterize the bicyclic graphs with nullity $n - 5$.

1 Introduction

In the 1980s, based on of Gustafson's theoretical proof, Fiedler and Markham first formulated the nullity of graphs theory. Later Markham and Fiedler will abstract theory into research matrix, makes this theory in real matrix and complex matrix has applications. Since then, domestic and foreign scholars of the nullity of graphs to chart makes a deep research, and obtained a series of significant conclusions.

The research for the nullity of graphs has a still olderchemical origin. Longuet–Higgins says a bipartite graph G (corresponding to a alternant hydrocarbons), if is singular, it means the molecule is unstable, this problem to the unbipartite graph (corresponding to the alternant hydrocarbons) also is significant. In recent years, relevant chart of the nullity of graphs research is very active, especially figure for some special graphs and the graph with less edges, still not getting good solution.

T. Zheng (✉)
School of Science, Anhui University of Science and Technology, Huainan 232001, China
e-mail: 109093273@qq.com

2 Preliminaries

Let $G = (V, E)$ be a simple undirected graph with vertex set $V = V(G) = \{v_1, v_2, \ldots, v_n\}$ and edge set $E = E(G) = \{e_1, e_2, \ldots, e_n\}$. For any $v \in V$, denote by $d(v)$ and $N(v)$ respectively the degree and neighborhood of v in G. A vertex $v \in V$ is called a pendant vertex if $d(v) = 1$, the vertex is called a quasi-pendant vertex if $w \in V$ adjacent to a pendant vertex of G. If W is a nonempty subset of $V(G)$, then the subgraph of G obtained by taking the vertices in W and joining those pairs of vertices in W which are joined in G is called the induced subgraph of G, denoted by $G[W]$. We write $G - \{v_1, v_2, \ldots, v_k\}$ for the graph obtained from G by removing the vertices $v_1, v_2, \ldots, v_k$ and all edges incident to any of them, we sometimes write $G - \{v_1, v_2, \ldots, v_k\}$ as $G - G_1$ if $G - \{v_1, v_2, \ldots, v_k\} = G_1$. The disjoint union of two graphs G_1 and G_2 is denoted by $G_1 \cup G_2$, with the vertex-set $V(G_1) \cup V(G_2)$ and the edge-set $E(G_1) \cup E(G_2)$. The graph with n vertices and no edges is the null graph of order n. Usually, the complete graph, cycle and star of order n are denoted by K_n, C_n and S_n. An isolated vertex is sometimes denoted by K_n. As usual, the star graph and the cycle of order n are denoted by $K_{1,n-1}$ and C_n, respectively. An isolated vertex is sometimes denoted by K_1.

The adjacency matrix $A(G)$ of graph G of order n, having vertex-set $V(G) = \{v_1, v_2, \ldots, v_k\}$ is the $n \times n$ symmetric matrix $[a_{ij}]$, such that $[a_{ij}] = 1$ if v_i and v_j are adjacent and 0, otherwise. A graph is said to be singular (non-singular) if its adjacency matrix A is a singular (non-singular) matrix. The eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of $A(G)$ are said to be the eigenvalues of the graph G, and to form the spectrum of this graph.

Definition 1.1 The number of zero eigenvalues in the spectrum of the graph G is called its nullity and is denoted by $\eta(G)$.

Let $r(A(G))$ be the rank of $A(G)$. Clearly, $\eta(G) = \eta(G) = n - r(A(G))$. The following result is clear.

Proposition 1.1 Let G be a graph with n vertices. Then $\eta(G) = n$ if and only if G is a null graph.

Proposition 1.2 (1) $\eta(k_n) = 0, \eta(k_{r,t}) = n - 2(r + t = n)$

(2) Let H be an induced subgraph of G, then $r(H) \leq r(G)$

(3) Let $G = G_1 \cup G_2$, then $r(G) = r(G_1) + r(G_2)$

(4) $\eta(C_n) = \begin{cases} 2, n \equiv 0 \pmod 4 \\ 0, n \neq 0 \pmod 4 \end{cases}$

It is well known that $0 \leq \eta(G) \leq n - 2$ if G is a simple graph on n vertices with at least one edge. In [1], Collatz and Sinogowitz first posed the problem of characterizing all graphs with $\eta(G) > 0$. This question is of great interest in chemistry, because, as has been shown in [2], for a bipartite graph G (corresponding to an alternant hydrocarbon), if $\eta(G) > 0$, then it indicates that the molecule which such a graph represents is unstable. The nullity of a graph is also important in mathematics, since it is related to the singularity of $A(G)$. The
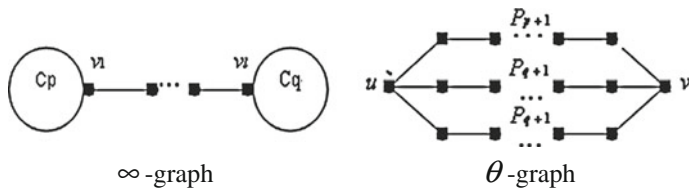
Fig. 1 ∞-graph and θ-graph

problem has not yet been solved completely. In recent years, this problem has been payed much attention by many researchers [2–8]. This paper consider the problem of bicyclic graphs with $\eta(G) = n - 5$.

A bicyclic graph is a simple connected graph in which the number of edges equals the number of vertices plus one. Let C_p and C_q be two vertex-disjoint cycles. Suppose that $v_1 \in C_p, vl \in C_q$. Joining v_1 and vl by a path $v_1 v_2, \ldots, v_n$ of length $l - 1$, where $l \geq 1$ and $l = 1$ means identifying $v1$ with vl, the resultant graph, denoted by $\infty(p, l, q)$, is called an ∞-graph. Let P_{l+1}, P_{p+1} and P_{q+1} be three vertex-disjoint paths, where $\min(p, l, q) \geq 1$ and at most one of them is 1. Identifying the three initial vertices and terminal vertices of them, respectively, the resultant graph, denoted by $\theta(p, l, q)$, is called a θ-graph. Then the bicyclic graphs can be partitioned into two classes: the class of graphs which contain an ∞-graph as an induced subgraph and the class of graphs which contain a θ-graph as an induced subgraph (Fig. 1).

Definition 1.2 Call a graph $\infty(p, l, q)$ (or $\theta(p, l, q)$) the base of the corresponding bicyclic graph G. Denote the base of G by $\chi(G)$.

3 Construct Bicyclic Graphs with $\eta(G) = n - 5$ and Main Results

The following lemmas are needed.

Lemma 2.1 *For a graph G containing a vertex of degree* 1, *if the induced subgraph H (of G) is obtained by deleting this vertex together with the vertex adjacent to it, then the relation $\eta(G) = \eta(H)$ holds* [7].
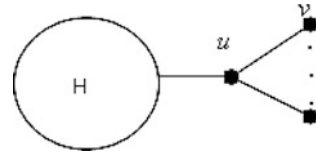
Lemma 2.2 *Let $G = G_1 \cup G_2 \cup \ldots \cup G_t$, and $G_1, G_2, \ldots, G_t$ are components of G. Then $\eta(G) = \sum_{i=1}^{t} \eta(G_i)$* [7].

Lemma 2.3 *Let G be a connected graph of order n. Then $r(G) = 2$ if and only if $G = k_{r,n-r}$; $r(G) = 3$ if and only if $G = k_{a,b,c}$, where $a + b + c = n$.*
Let H is he induced subgraph of G, we begin to discuss two cases [5].

Case 1 $n_H = 3, G$ contains pendant vertices.

Fig. 2 The graph with
$n_H = 3$

As shown in the Fig. 2, $\eta(G) = \eta(G - u - v) = n - 5, r(A(G)) = 5, r(A(G - u - v)) = 3$ then we need to discuss the case of $r(H) = 3$. $r(H) = 3$ if and only if $H = k_{a,b,c}(a + b + c = n_H)$.

Subcase 1.1 If n_H=3, the bicyclic graphs as shown in Fig. 3.

Subcase 1.2 If $n_H = 4$, the bicyclic graphs as shown in Fig. 4.

Case 2 The bicyclic graphs without pendant vertice can be partitioned into two classes: ∞-graph and θ-graph.

Lemma 2.4 *The bicyclic graphs with rank 5 are* $\infty(3, 1, 3)$ *and* $\infty(3, 1, 4)$ *(It also contains* $\infty(4, 1, 3)$*, it is no need repeat, because the graph is undirected graph and isomorphic) or* $\theta(2, 2, 3)$*.*

Proof The graphs with rank 5 cannot contain $C_3 \cup C_3, P_6$ or $C_3 \cup P_3$ as an induced subgraph, since $r(C_3 \cup C_3) = 6, r(P_6) = 6, r(C_3 \cup P_3) = 6$.

Case 1 Let $\chi(G) = \infty(p, l, q)$, if $p = 3$, $q = 3$, $l \geq 2$ then $C_3 \cup C_3$ is an induced subgraph of G, it is impossible.
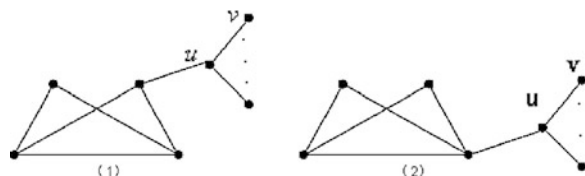
If $p = 3$, $q \geq 4$ or $p \geq 4$, $q = 3$, $l \geq 2$ and $p \geq 4$, $q \geq 4$, $l \geq 1$ then $C_3 \cup P_3$ or P_6 is an induced subgraph of G it is impossible. It shows that $r(\infty(3, 1, 3)) = 5$ and $r(\infty(3, 1, 4)) = 5$ (Fig. 5).

Case 2 $\chi(G) = \theta(p, l, q)$. In this case, the length of the induced cycle must be less or equal to 5 (Fig. 6).

Fig. 3 K(1, 1, 1)
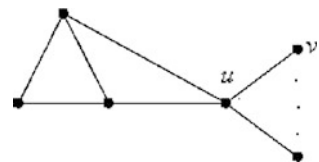
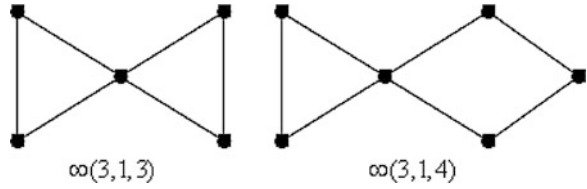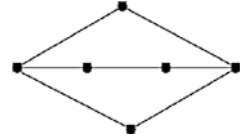

Fig. 4 The graph with
$n_H = 4$

(1)   (2)

Fig. 5 $\infty(3,1,3)$ and $\infty(3,1,4)$

$\infty(3,1,3)$ $\infty(3,1,4)$
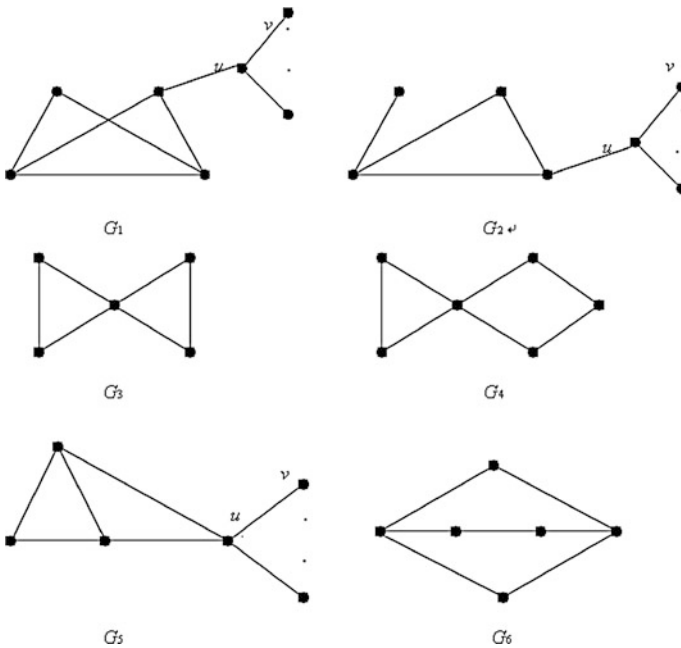
Fig. 6 $\theta(2,2,3)$

    Among these graphs, only $r(\theta(2,2,3))$.

    Through the above certificate and relevant conclusion, we can characterizes bicyclic graphs with nullity $n-5$.

Definition 2.1 Let G is a bicyclic graph, if and only if $G = G_i$ ($1 \le i \le 6$) (Fig. 7), $\eta(G) = n - 5$.

G_1 G_2

G_3 G_4

G_5 G_6

Fig. 7 Bicyclic graphs with nullity $n-5$

4 Conclusion

At present about figure the nullity, many scholars had in-depth research, but still remains the singularity problems and the singular figure depicting problems still not getting good solution. The nullity of graphs research has wider range. This paper characterizes bicyclic graphs with nullity $n-5$ We solve the problem from three parts: (1) The research situation of the nullity of bicyclic graphs. (2) Construct bicyclic graphs with $\eta(G) = n - 5$. (3) Through the analysis and certificate we can characterizes bicyclic graphs with nullity $n - 5$.

For the nullity of graphs research has wider range, such as the research of trees, the unicyclic graphs, the bicyclic graphs and the linear graphs. In the future, we will do a lot of works to show the power of the nullity theorem and the related theorems, we apply them to semiseparable and related matrices.

References

1. Collatz L, Sinogowitz U (1957) Spektren endlicher Grafen. Abh. Math Univ Hamburg, 2163–2167
2. Hu S, Xuezhong T, Liu B (2008) On the nullity of bicyclic graphs. Linear Algebra Appl, 4291387–4291391
3. Sciriha I (1999) On the rank of graphs. In: Alavi Y, Lick DR, Schwenk A (eds) Combinatorics, graph theory and algorithms, vol 2. Michigan, pp 769–778
4. Sciriha I (1998) On the construction of nullity one. Discrete Math 181193–181211
5. Sciriha I (1998) On singular line graphs of trees. Congressus Numeratium, 13573–13591
6. Tan XZ, Liu BL (2005) On the nullity of unicyclic graphs. Linear Algebra Appl 408:212–222
7. Cvetkovic D, Doob M, Sachs H (1980) Spectra of graphs. Academic Press, New York
8. Gong S-C, Fan Y-Z, Yin Z-X (2009) The tree with second largest Laplacian spread. Academic Press, Hangzhou

Study on Security Supervising and Managing Methods of the Trusted Cloud Computing Based on Petri Net

Lu Liu, Xian-wen Fang, Xiang-wei Liu and Jing Ji

Abstract With the constant popularization of cloud computing, the importance of security problem is gradually increasing and becomes the important factor which restricts cloud computing development. Currently, combination of cloud computing and trusted computing technology are the main topic on cloud security. In the paper, according to analyzing the research situation of the cloud computing security, a novel trusted cloud computing technology according to the related technology is proposed, and a secure supervising and managing methods by behavioral analysis and service composition is analyzed based on Petri net.

Keywords Cloud computing · Trust computing · Petri net · Secure

1 Introduction

Currently, cloud computing is one of the hotspot issues in industry and academia. Cloud computing was firstly proposed by the IBM company in the Technical White Paper of 2007, and became a hot topic in IT from then on. The idea of cloud computing is the commercial implementation of public computer resource and

L. Liu (✉) · X. Fang · X. Liu · J. Ji
Department of Information and Computing Science, Anhui University of Science and Technology, Huainan 232001, China
e-mail: luliu@hotmail.com

X. Fang
e-mail: fangxianwen@hotmail.com

X. Liu
e-mail: xwliu@hotmail.com

J. Ji
e-mail: jingji@hotmail.com

provides a new idea for information management and service [1]. However, at present, with the spread of cloud computing, the importance of security is gradually rising, and becomes an important factor to restrict its development. Thus, attracting the enterprise and the organization to apply cloud computing technology and platform on a large scale, as well as giving them secure business administration for their data in cloud service, we must analyze roundly and set about solving all kinds of problem that security cloud computing faced.

In theoretical research, the currently primary researches at home and abroad in cloud computing security included: (1) a resource access control scheme across multi-domains [2]. (2) Security access control and data validation, because the user-defined access control by service provider was untrusted, researchers cared about how to implement the access control of data in cloud computing model [3]. (3) Virtual security technology, this virtual technology is key of core technology to realize cloud computing, and cloud frame provider of cloud computing platform by virtual technology had to guarantee security of users and isolation. (4) Private data protection [4], with the deepening research of cloud computing security, trusted computing is proposed to solve cloud computing security problems. Thus, in practice, trusted computing is a key problem of cloud computing security researches.

Through analysis of the security problem in the current cloud computing and influence of cloud computing on the field of information security, this paper put forward the relatively targeted suggestions to promote the development of the future cloud computing industry in Anhui province of China. At the same time, the discovery and elimination technology of untrusted problems as well as the criterion of content and technological specification is analyzed.

2 The Security Analysis of Cloud Computing

Cloud computing takes dynamic calculation of services as the main technical characteristics and the agile "service contract" as the core business characteristics. This model brings a huge impact into the field of information security. Cloud computing is built on the internet, thus network structure and equipment composition determine that the security problems of traditional network on cloud computing will still remain, such as network interruption caused by Viruses, Trojan invasion, data leakage, hardware and software, etc. In addition, cloud computing integrates with many business functions. Its advantages are wide geographical coverage, carrying many business, transmitting information rapidly, mass data storage, and so on. These characteristics make the cloud computing platform about the network and information to face with the security challenges which are more severe than the traditional Internet service. For example, the uncontrollable problems are caused by excessive dependence on external resources, and unavailable data center is caused by external forces which network suffered, even the protection problem of mass user data. The main root is reflected in: (1) all

kinds of cloud application operated on cloud platform have no fixed infrastructure and secure border, so it's hard to realize the security of user data and privacy protection; (2) since resources which cloud services involve are held by many managers, causing interest conflicts, it's impossible to unitedly plan and deploy the security measures; (3) in cloud platform, data and calculation is highly centralized, so security measures must satisfy the needs of mass information processed [5].

Therefore, by analyzing the recent research situation, we find that cloud computing technology has more than the security problems of traditional network. For diverse business, the cloud computing platform exists the security problem. These problems are not only reflected in management level (including standards, regulations and supervising system, etc.), it also needs to research the related management systems from the technical level, to form the adapted countermeasures of a new technology system and management system (standards, regulations and supervision system) gradually, then we can realize the target about the safety and credibility of cloud computing technology.

3 Security Management System of Trusted Cloud Computing Based on Petri Net

Cloud computing brings great benefits while brings huge destructive capability to people. So in the development of cloud computing, the technology of supervising system should be vigorous expansion, and firmly grasp initiative of technology to prevent controlment and utilization by malicious persons. Compared with supervising and managing system of the Internet, due to characteristics of cloud computing, it is more difficult to realize real-time defense and online supervising. On the one hand, the supervising model of operation in cloud computing platform is analyzed for real-time analysis and online supervising, to realize the fast discrimination of security attack, pre warning and protection. On the other hand, the content of cloud computing needs to be supervised by using relative technology for filtering content, as well as establish standardized criteria of business and content, to realize effective supervision of cloud computing content.

3.1 The Research of Trust Computing Security

Cloud computing is unifying scheduling and organization through the cloud computing. Cloud computing platform provides variety of shared resources as well as free application program and memory space. Since cloud computing platform puts the application program and stored material of customers in cloud, security of cloud computing is the key issue to management of cloud computing platform. However, cloud computing platform is not completely believable. Once one or

some nodes of cloud capture by hackers, the hacker may continue to attack and control cloud computing platform through the nodes, and in the end control supercomputing capabilities cloud computing platform organized, to implement greater destruction and attack of the network. Consequently, the research of cloud computing is a significant trend for secure development of cloud computing.

With combining the technology of trust computing with cloud computing, to provide cloud services in a reliable way becomes a hot topic of the research field of cloud security. Santos et al. presents a trusted cloud computing platform (TCCP) model, and IaaS service providers can give an airtight execution environment of the box-type based on this platform [6]. Moreover, it allows users to examine the security of the service before launch virtual machine. Sadeghi et al. think the technology of cloud computing provides trusted software and hardware, as well as proves credible mechanism of their own behavior [7]. Meanwhile it can be used to resolve confidentiality and integrity problem of outsourcing data. They design a kind of credible software token, which bounds to each other with the verification module of secure function, to execute various functional operation for outsourcing sensitive (encrypted) data.

However, the traditional trust computing can't only rapidly discover the untrusted problem. We present a novel trusted cloud computing technology based on Petri net. Petri net has two advantages: one is in analyzing models, consistency between models from a behavioral angle is considered. Another is in finding change region, speediness is realized [8]. In other word, technology of Petri net can realize real-time monitoring for change nodes in cloud computing. In consequence, on the one hand, the way of finding change region based on Petri net can real-timely monitor the change nodes in trusted cloud computing platform. According to Monika et al. [9], we can locate the region of change propagation in models by the change node of the target model and then real-timely monitor it. Once the untrusted change region is found, it will be deleted or separated for preventing its propagation. Otherwise, the region will be marked as the trusted region and monitor the next change node. By analogy, this technology can timely check to every new node, thus strengthen security management of the cloud computing platform. On the other hand, the way of bisimulation between models based on Petri net can guarantee behavioral consistency between similar models to maintain safe operation with customers, thus strengthen security prewarning and protection.

3.2 Security Supervising and Managing System of the Trusted Cloud Computing

Cloud computing integrates many business functions. So its advantages are the wide coverage, the large range of application, much business, rapid propagation of information, the large amount of data storage and so on. This means that cloud
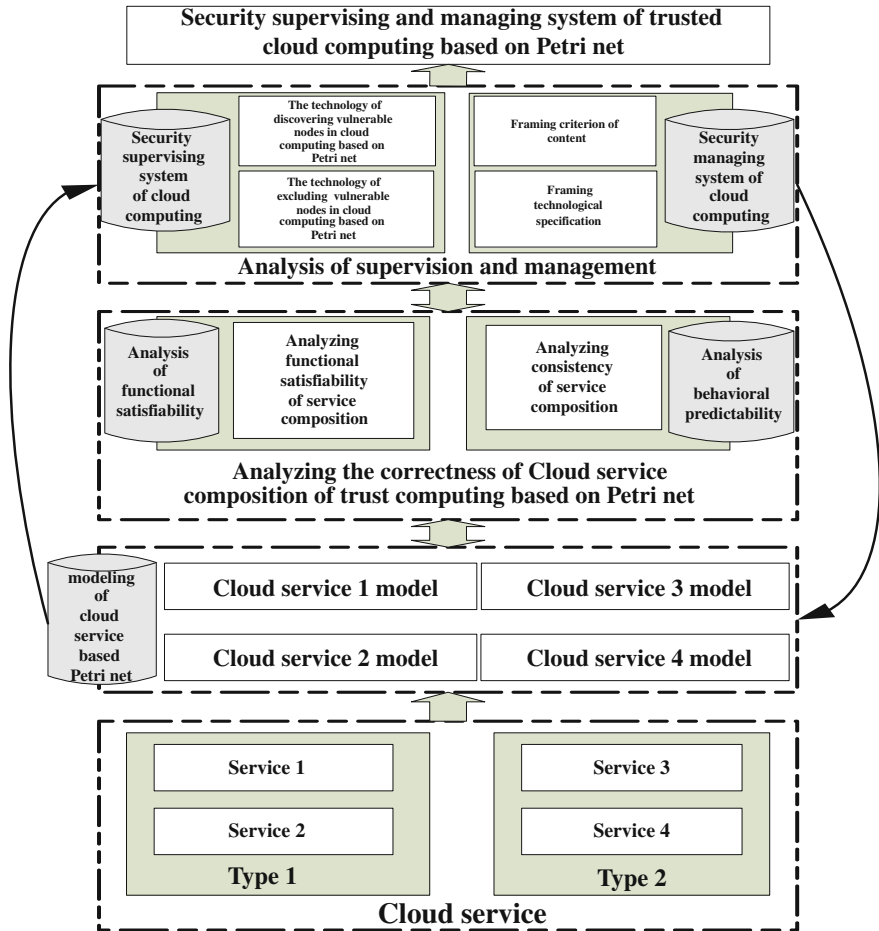
Fig. 1 Security supervising and managing system

computing platform will face the more severe security issue. Today, because locating business of cloud computing is sill indistinct, there is difficult to classify cloud computing service as a certain type of business. At the same time, development of cloud computing services is irregularity. And it is difficult to build the sound supervising and managing system. So building the security supervising and managing system is another key issue in development of cloud computing. We solve this problem from the following points.

Firstly, confirming the goal of trusted cloud computing security supervising and the requirement of cloud computing security management is the core content. The goal is to guarantee security and trust of cloud computing. In order to realize this goal, we merge trust computing with security management, which is presented in the fourth layer of Fig. 1. The requirement is in form of government function and

laws for the method, to guarantee security, authenticity and openness of cloud services. In order to realize this requirement, we frame the corresponding criterion of content and technological specification. Figure 1 satisfies the protection of business and content in cloud computing by the analyzing technology of security management, pre warning and processing of illegal business and content, and the specification of optional models and preserving service log.

Secondly, cloud service composition is analyzed. We use the technology of service composition based on Petri net in Fig. 1. This technology composes models to reducing the number of models as well as convenient management and supervision. At the same time, we analyze functional satisfiability and behavioral predictability by related technology based on Petri net, to guarantee the original purpose of models after composition.

Finally, cloud computing is real-timely supervised and managed. In order to prevent burglary and peep, Fig. 1 show that once untrusted nodes are found, this system can repair and restructure after returning the previous link, even being deleted or separated. Thus, it can effectively prevent the propagation of untrusted nodes and narrow the range of its hazard to strengthen security prewarning and protection.

Based on the above conclusions, we establish the supervising and managing system of trust computing like Fig. 1. It can not only realize the real-time supervision of different business and content in cloud computing, but also rapidly find and solve change nodes.

4 Conclusion

With the development of IT industry, the application of cloud computing is increasingly extension. So the security of cloud computing becomes the urgent problem. This paper establishes a new security supervising and managing system based on Petri net, combining cloud computing and trusted cloud computing. We solve the secure problem from three parts: (1) the research situation of the cloud computing security is analyzed. (2) Trusted cloud computing is researched by some related technology based on Petri net. (3) The security supervising and managing system of trusted cloud computing is given from the angle of behavior and service composition based on Petri net.

However, the factors of threatening security of cloud computing are diverse. It is not enough to only consider the technology. To solve this problem need joint efforts by information security academia, industry and government.

In the future, we will analyze the secure system of object-oriented cloud computing based on Petri net. And then construct the lager security supervising and managing system of cloud computing.

References

1. Liu Q (2011) The research of cloud computing services problem. Internet Inf Secur 6:36–37
2. Shafiq B, Joshi JBD, Bertino E, Ghafoor A (2005) Secure interoperation in a multidomain environment employing RBAC policies. IEEE Trans Knowl Data Eng 17(11):1557–1577
3. Hong C, Zhang M, Feng DG (2010) A way of ciphertext-access control in cloud storage. Compu Res Dev 47:259–265
4. Roy I, Ramadan HE, Setty STV, Kilzer A, Shmatikov V, Witchel E, Airavat (2010) Security and privacy for map reduce. In: Proceedings of the 7th Usenix Symposium on networked systems design and implementation, 297–312
5. Feng DG, Zhang M, Zhang Y, Xu ZH (2011) Research of cloud computing security. J Softw 22(1):71–83
6. Santos N, Gummadi KP, Rodrigues R (2009) Towards trusted cloud computing. In: Proceedings of the workshop on hot topics in cloud computing 2009, San Diego
7. Sadeghi AR, Schneider T, Winandy M (2010) Token-based cloud computing: secure outsourcing of data and arbitrary computations with lower latency. In: Proceedings of the 3rd International Conference on Trust and Trustworthy Computing, 417–429
8. Matthias W, Jan M, Mathias W (2012) Propagating Changes between aligned process models. J Syst Softw, 1885–1898
9. Monika W, Modood A, Falko K, Frank L, Thomas R, David S (2011) Business process change management based on process model synchronization of multiple abstraction levels. In: 2011 IEEE international conference on service-oriented computing and applications, 1–4

A Method for Security Evaluation in Cloud Computing Based on Petri Behavioral Profiles

Xianwen Fang, Mimi Wang and Shenbing Wu

Abstract To solve the problem of security evaluation in cloud computing, the modeling and analyzing methods of security evaluation about cloud computing are very important. Firstly, extending behavioral profiles into comprehensive behavioral profiles. Then, studying the degree of keeping consistent behavior relation which looks as the objective evaluation function based on comprehensive behavioral profiles. Finally, using behavioral theory of Petri Net to solve the evaluation function, a new method for security evaluation in cloud computing is proposed. The theoretical analysis and specific example show that the method is very effective.

Keywords Cloud computing · Petri net · Comprehensive behavioral profiles · Security evaluation

1 Introduction

Recently the cloud computing paradigm has been receiving significant excitement and attention in the media and blogosphere. To some, cloud computing seems to be little more than a marketing umbrella, encompassing topics such as distributed computing, grid computing, utility computing, and software-as-a-service, that have

X. Fang (✉) · M. Wang · S. Wu
Department of Information and Computing Science, Anhui University of Science
and Technology, Huainan 232001, China
e-mail: fangxianwen@hotmail.com

M. Wang
e-mail: weiluwangmimi@hotmail.com

S. Wu
e-mail: shbwu@hotmail.com

already received significant research focus and commercial implementation [1]. Naturally, we realize the significant role the security problem plays in cloud computing is attracted importance to with the development of cloud computing, so how to solve the problem of security evaluation in cloud computing is becoming more and more important.

For security evaluation in cloud computing, there were some previous studies, but the majority is not on the basis of process behavior [2–8]. Ristov et al. [2] gave an overview of information security standards and evaluated their completeness, proposed a new extension to the ISO 27001:2005 standard including a new control objective about virtualization applicable for cloud systems. Dawei et al. [3] gave the definition of trust in cloud systems, analyzed the properties of trust by referring to the fruits from social science, proposed a dynamic multi-dimensional trust model named DMTC, and gave an algorithm based on the DMTC model to compute trust in cloud systems. Wei et al. [4] proposed a trust mechanism-based task scheduling model for cloud computing, built trust relationship among computing nodes referring to the trust relationship models of social persons, evaluated the trustworthiness of nodes by utilizing the Bayesian cognitive method, and proposed the trust dynamic level scheduling algorithm for cloud computing. Qiang et al. [5] introduced the definition of trust in cloud systems, analyzed the properties of trust, proposed an extensible trust evaluation model named ETEC, and gave an algorithm based on the ETEC model to compute trust in cloud systems. Wenchuan et al. [6] proposed and solved the privacy-preserving fuzzy comprehensive evaluation problem in the area of secure multi-party computation, and analyzed the protocol's security and privacy-preserving protocol of finding the maximum component's position. Hyukho et al. [7] presented a trust model for efficient reconfiguration and allocation of computing resources satisfying various user requests, prepared the best available resources for each service request in advance, providing the best resources to users, and carried out experiments for reliability analysis with 4 data types, including an all data set, random data set, recent data set, and the data set within a standard deviation. Qintian et al. [8] discussed evaluation importance of user behavior trust and evaluation strategy, in the cloud computing, including trust object analysis, principle on evaluating user behavior trust, basic idea of evaluating user behavior trust, evaluation strategy of behavior trust for each access, and long access.

For process behavior, previously there were some researches [9–13]. Such as we presented previously a behavior trustworthiness analysis method of business process based on induction information, analyzed the consistent behavior relativity to guarantee the predictable function, proposed a process mining methods based on induction information for the external factors to analyze the behavior change of business process, and gave out experiment simulation [9]. Matthias and Jan et al. proposed the concept of behavioral profiles that captured the essential behavioral constraints of a process model, showed that these profiles could be computed efficiently, used behavioral profiles for the definition of a formal notion of consistency, which was less sensitive to model projections than common criteria of behavioral equivalence and allowed for quantifying deviation in a metric way, and

then implemented the derivation of behavioral profiles and the calculation of a degree of consistency [10–13], but they did not research indirect behavior relation.

Based on the above background, our approach starts from a motive example, leads to the necessity of existing the indirect behavior relation in daily life, which draws forth the usefulness of studying indirect behavior relation. By studying the security evaluation method based on extended behavior profile of Petri nets, this paper starting from the point of view of indirect behavior relation, research consistency behavior keep of Petri nets, and measure the degree of consistency behavior keep, quantifying the consistency behavior keep of the Petri net.

The remainder is organized as follows. Section 2 gives a method for looking for the objective evaluation model and an example as the basis of the study. In Sect. 3, we provide a kind of method based on the comprehensive behavioral profiles to analyze the problem of security evaluation in cloud computing. Conclusions are in Sect. 4.

2 Objective Evaluation Model About the Cloud Security

2.1 Basic Concepts

This section describes the basic concepts and definitions used in our approach, notion of Petri net and weak order can reference literature [14]. Below we mainly introduce the concept of comprehensive behavioral profiles.

Definition 1 (*Comprehensive Behavioral Profile*) Let (N, M_0) be a net system with $N = (P, T, F)$. A pair of transitions $(x, y) \in (T \times T)$, $(y, z) \in (T \times T)$ can be in the following profile relations:

1. The simulated strict order relation $x \Longrightarrow z$, iff $x \succ y$, $y \not\succ x$, $y \succ z$ and $z \not\succ y$.
2. The simulated exclusiveness relation $x \overline{+} z$, iff $x \not\succ y$, $y \not\succ x$, $y \succ z$ and $z \not\succ y$.
3. The simulated interleaving relation $x \overline{\|} z$, iff $x \succ y$, $y \succ x$, $y \succ z$ and $z \not\succ y$.

$B_1 = \{\rightarrow, +, \|, \Longrightarrow, \overline{+}, \overline{\|}\}$ is the comprehensive behavioral profile over T.

We give the notion of transition pairs of keeping consistent behavior relation as follows.

Definition 2 (*Transition Pairs of Keeping Consistent Behavior Relation*) Let $PN_1 = (P_1, T_1, F_1, M_{01})$, $PN = (P, T, F, M_0)$ be two live and bounded Petri nets and $B_1 = \{\rightarrow_1, +_1, \|_1, \Longrightarrow_1, \overline{+}_1, \overline{\|}_1\}$, $B = \{\rightarrow, +, \|, \Longrightarrow, \overline{+}, \overline{\|}\}$ be their comprehensive behavioral profiles, respectively. Let $Re_1 \in B_1 \cup \{\Longrightarrow_1^{-1}\}$, $Re \in B \cup \{\Longrightarrow^{-1}\}$ $T_1' = T_1 \cap T$. The set of comprehensive behavioral profile transition pairs of keeping consistent behavior relation t_{12} for PN_1 contains all pairs $(t_x, t_y) \in T_1' \times T_1'$, such that

1. if $t_x = t_y$, then $\forall t_s \in T$ it holds, $(t_x \mathrm{Re}_1 t_x^\wedge t_s \mathrm{Re} t_s) \Rightarrow \mathrm{Re}_1(PN_1) \cong \mathrm{Re}(PN)$
2. if $t_x \neq t_y$, then $\forall t_s, t_t \in T$ it holds $(t_x \mathrm{Re}_1 t_y \wedge t_s \mathrm{Re} t_t) \Rightarrow \mathrm{Re}_1(PN_1) \cong \mathrm{Re}(PN)$.

The set $\mathrm{Re}^{\cong} T'$ for PN_2 is defined analogously.

2.2 Objective Evaluation Model of the Cloud Security

Figure 1 shows the integrating Petri net model of cloud services systems in cloud computing environment.

In the Fig. 1, we can see that $t_{16} \overline{+} t_{19}$, $t_{16} \overline{\rightarrow} t_{20}$, $t_{16} \rightarrow t_{23}$, $t_6 \| t_{21}$, $t_{15} \rightarrow t_{21}$, $st_{18} \| t_{21}$, moreover, the transition t_{16} and t_6 seem to be high-risk place of safety loophole. To attain the most safety, it must be that t_{16} and t_{19} would be best in the simulated interleaving relation, t_{16} and t_{20} would be best in the strict order relation, t_{16} and t_{23} would be best in the exclusiveness relation, t_{16} and t_{21} would be best in the simulated interleaving relation, t_{15} and t_{21} would be best in the simulated strict relation, t_{18} and t_{21} would be best in the strict order relation, that is to say $t_{16} \| t_{19}$, $t_{16} \rightarrow t_{20}$, $t_{16} + t_{23}$, $t_{16} \| t_{21}$, $t_{15} \overline{\rightarrow} t_{21}$, $t_{18} \rightarrow t_{21}$. So we can obtain the objective evaluation model such as Fig. 2.
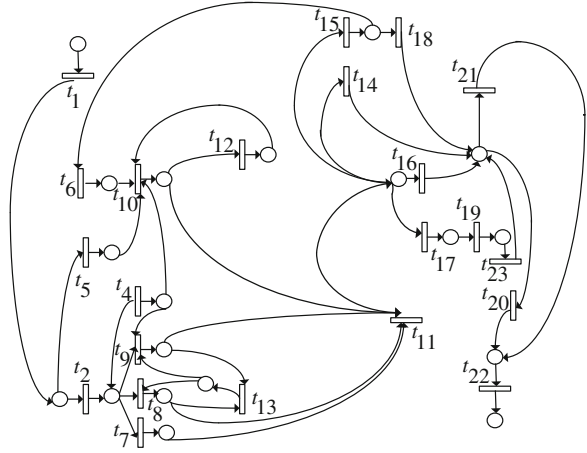
3 Objective Evaluation Function About the Cloud Security

Below we use behavioral theory method to establish the objective function, firstly study the behavior relationship between transitions of Figs. 1 and 2, then obtain the object evaluation function—degree of keeping consistent behavior relation

Fig. 1 A Petri net model of cloud services systems
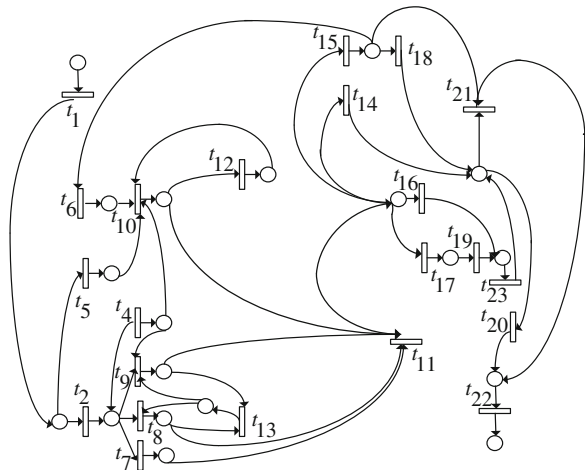
Fig. 2 Petri net of target
evaluation model

based on comprehensive behavioral profile, and last use an algorithm to calculate the degree. The objective evaluation function shows such as formula (1).

Objective evaluation function:

$$\mathrm{Re}_D = \frac{\omega_1 \left| \mathrm{Re}^{\cong} T_1' \right| + \omega_2 |\mathrm{Re}^{\cong} T'|}{\left| (T_1' \times T_1') \right| + |(T' \times T')|} + \frac{(1 - \omega_1) \left| \mathrm{Re}^{\sim} T_1' \right| + (1 - \omega_2)|\mathrm{Re}^{\sim} T'|}{\left| (T_1' \times T_1') \right| + |(T' \times T')|} \quad (1)$$

(where ω_1, ω_2 is the ratio of indirect weak order relation and the overall weak order relation in $T_1' \times T_1'$ and $T' \times T'$, respectively.)

We call Re_D as the degree of keeping consistent behavior relation based on comprehensive behavioral profile. The following is an algorithm of calculating the Re_D.

Algorithm 1 *Calculation for* Re_D

Input: two live and bounded Petri nets $PN_1 = (P_1, T_1, F_1, M_{01})$, $PN = (P, T, F, M_0)$, sets of transitions $T_1 = \{t_{11}, t_{12}, \cdots, t_{1n}\}$, $T = \{t_{21}, t_{22}, \cdots, t_{2m}\}$

Output: Re_D

(1) Analyze $t_{11}, t_{12}, \cdots t_{1n}$ in T_1, take t_i, t_j for example, if $t_i \succ t_j$ and also meets $t_i \succ t_j$ in T, output $\mathrm{Re}^{\sim} T_1' = T_1' \times \mathrm{T}_1'$, and then return to step (2); otherwise, remove (t_i, t_j) from $T_1' \times T_1'$, output $\mathrm{Re}^{\sim} T_1' = (T_1' \times T_1') \backslash \{(t_i, t_j)\}$, then return to step (2).

(2) Similarly, implement above steps, until the last two transitions t_{1n-1}, t_{1n} in T_1, if $t_{1n-1} \succ t_{1n}$, and also meets $t_{1n-1} \succ t_{1n}$ in T, output $\mathrm{Re}^{\sim} T_1' = T_1' \times \mathrm{T}_1'$, and then return to step (3); otherwise, remove (t_{1n-1}, t_{1n}) from $T_1' \times T_1'$, output $\mathrm{Re}^{\sim} T_1' = (T_1' \times T_1') \backslash \{(t_{in-1}, t_n)\}$, then return to step (3).

(3) Analyze $t_{1i}, t_{1i+1}, \cdots, t_{1j}$ in $\mathrm{Re}^{\sim} T_1$, take $t_{1i}, t_{1i+1}, t_{1i+2}$ for example, if it meets $t_{1i} \overline{\Longrightarrow} t_{1i+2}$ in PN_1, and also in PN, output $\mathrm{Re}^{\cong} T_1' = \mathrm{Re}^{\sim} T_1'$, and then return to step (4); otherwise, return to step (4).

(4) If it meets $t_{1i} \mp t_{1i+2}$ in PN_1, and also in PN, output $\mathrm{Re}^{\cong}T_1' = \mathrm{Re}^{\sim}T_1'$, and then return to step (5); otherwise, return to step (5).

(5) If it meets $t_{1i}\|t_{1i+2}$ in PN_1, and also in PN, output $\mathrm{Re}^{\cong}T_1' = \mathrm{Re}^{\sim}T_1'$, and then return to step (5); otherwise, output $\mathrm{Re} \cong T_1' = (T_1' \times T_1')\backslash\{(t_{1i}, t_{1i+2})\}$, and then return to step (6).

(6) Similarly, implement above steps, until the last three transitions $t_{1j-2}, t_{1j-1}, t_{1j}$ in $\mathrm{Re}^{\sim}T_1$, if it meets $t_{1j-2} \longrightarrow t_{1j}$ in PN_1, and also in PN, output $\mathrm{Re}^{\cong}T_1' = \mathrm{Re}^{\sim}T_1'$, return to step (9); Otherwise, return to step (7).

(7) If it meets $t_{1j-2} \mp t_{1j}$ in PN_1, and also in PN, output $\mathrm{Re}^{\cong}T_1' = \mathrm{Re}^{\sim}T_1'$, return to step (9); otherwise, return to step (8).

(8) If it meets $t_{1j-2}\|t_{1j}$, in PN_1, and also in PN, output $\mathrm{Re}^{\cong}T_1' = \mathrm{Re}^{\sim}T_1'$, return to step (9); Otherwise, output $\mathrm{Re}^{\cong}T_1' = (T_1' \times T_1')\backslash\{(t_{1j-2}, t_{1j})\}$, return to step (9).

(9) Similarly, implement above steps, find $\mathrm{Re}^{\cong}T'$, return to step (10).

(10) By formula (1), calculate Re_D, output Re_D, algorithm terminates.

According to the algorithms 1 and definition 1, we calculate the degree of keeping consistent behavior relation based on comprehensive behavioral profile of Figs. 1 and 2. As $|T_1' \times T_1'| = 23 \times 23 = 529$, $|T' \times T'| = 23 \times 23 = 529$; and according to the Algorithm 1, it holds $|\mathrm{Re}^{\sim}T_1'| = 529$, $|\mathrm{Re}^{\sim}T'| = 529$, $|\mathrm{Re}^{\cong}T_1'| = 17$ and $|\mathrm{Re}^{\cong}T'| = 17$; according to definition 1, it holds $\omega_1 = \frac{17}{529+17}$, $\omega_2 = \frac{6}{529+17}$ then the degree of keeping consistent behavior relation based on comprehensive behavioral profile is $\mathrm{Re}_D \approx 0.979615$.

By the above, we can see that the degree of keeping consistent behavior relation based on comprehensive behavioral profile is about 0.979615, it is close to 1, but less than 1, which shows only about 97.9615 % keep consistent behavior relation, about 2.0385 % of behavior relation is behavior inconsistent. That is to say, 97.9615 % of behavior relation is safe while 2.3085 % of behavior relation is not safe. How to ensure 2.0385 % be behavior consistent will be our future research work.

4 Summaries

In this paper, on the basis of previous study, we extend the behavioral profile, study the behavior keep of comprehensive behavioral profile, and propose a method to analyze the consistent behavior between the original model and the target model, use behavioral theory of the Petri net to establish the objective evaluation function of cloud security. On the basis of the evaluation function, an algorithm about the degree of keeping consistent behavior relation based on comprehensive behavioral profile is presented. Using the proposed-above method, the cloud service system in the cloud computing environment is analyzed.

References

1. Daniel JA (2009) Data management in the cloud: limitations and opportunities. IEEE Data Eng Bull 32:3–12
2. Ristov S, Gusev M, Kostoska M (2012) A new methodology for security evaluation in cloud computing. In: Ristov S, Gusev M, Kostoska M 35th international convention on information and communication technology, electronics and microelectronics. IEEE, Piscataway, NJ, USA, pp 1484–1489
3. Dawei S, Guiran C, Lina S, Fengyun L, Xingwei W (2011) A dynamic multi-dimensional trust evaluation model to enhance security of cloud computing environments. Int J Innovative Comput Appl 3:200–212
4. Wei W, Guosun Z, Junqi Z, Daizhong T (2012) Dynamic trust evaluation and scheduling framework for cloud computing. Secur Commun Netw 5:311–318
5. Qiang G, Dawei S, Guiran, C, Lina S, Xingwei W (2011) Modeling and evaluation of trust in cloud computing environments. In: Proceedings of the 2011 3rd international conference on advanced computer control (ICACC 2011). doi:10.1109/ICACC.2011.6016378
6. Wenchuan Y, Bei J, Bowei C (2012) Study of a fuzzy comprehensive evaluation problem in cloud computing. Artif Intell Comput Intell. doi:10.1007/978-3-642-33478-8_16
7. Hyukho K, Hana L, Woongsup K, Yangwoo K (2010) A trust evaluation model for QoS guarantee in Cloud systems. Grid Distrib Comput 3:1–10
8. Qintian L, Chuang L, Ni Y (2010) Evaluation of user behavior trust in cloud computing. In: 2010 international conference on computer application and system modeling. doi:10.1109/ICCASM.2010.5620636
9. Fang XW, Jiang CJ, Yin ZX (2011) The trustworthiness analyzing of interacting business process based on the induction information. Comput Sci Inf Syst 8:843–867
10. Matthias W, Jan M (2012) Perceived consistency between process models. Inf Syst 37:80–98
11. Matthias W, Artem P, Nirmit DI, Jan M, Mathias W (2011) Process compliance analysis based on behavioural profiles. Inf Syst 36:1009–1025
12. Matthias W, Holger Z, Jan M (2011) Optimizing complex event queries over business processes using behavioural Profiles. In: Matthias W, Holger Z, Jan M, ZurMuehlen M, Su JW Lecture notes in business information processing. Springer, Heidelberg, pp 743–754
13. Matthias W, Jan M, Mathias W (2011) Efficient consistency measurement based on behavioral profiles of process models. IEEE Trans Softw Eng 37:3410–3429
14. Matthias W (2011) Behavioural profiles—a relational approach to behaviour consistency. Institutional Repository of the University of Potsdam. http://opus.kobv.de/ubp/volltexte/2011/5559/URNurn:nbn:de:kobv:517-opus-55590

The Homomorphic Encryption Scheme of Security Obfuscation

Gao-Xiang Gong, Zheng Yuan and Xiao Feng

Abstract The cloud storage service, according to the data on the cloud computing safety protection problem, the paper presents secure obfuscating homomorphism encryption scheme. Constructing a point function obfuscator that based on perfectly one way probability hash function in scheme, construction depends on hash function and the computational difficulty problems, then use the computational difficulty problems, to realize the encrypted homomorphism function, also guarantee the function of the point function obfuscator at the same time, the scheme raises the security of the encrypted data. This paper provides the security proof of the scheme, and shows that the scheme is feasible.

Keywords Obfuscation · Computational difficulty · Point function obfuscation · Homomorphic encryption · Perfect one-way hash function

1 Introduction

The idea of Fully Homomorphic Encryption (FHE) was proposed by Rivest in 1978 [1]. The existence construction has long been one of the important open questions in cryptography. Unfortunately, there has been no progress in cryptographic circle after the issue was put forward for more than 30 years. The construction of fully homomorphic encryption problem research has achieved significant breakthrough in 2009. Gentry constructed the first fully homomorphic

G.-X. Gong (✉) · X. Feng
School of Telecommunications Engineering, Xidian University, Shaanxi 710071, China
e-mail: longyuggx@163.com

Z. Yuan
Beijing Electronic Science and Technology Institute, Beijing 100070, China
e-mail: zyuan@mail.tsinghua.edu.cn

encryption scheme [2] based on ideal lattice in 2009. The security of the scheme is based on the computational difficulty of two problems: Bounded distance coding issues over ideal lattice and spare subset sum problem. More detail about construction of the scheme is given in Gentry's dissertation [3]. A very simple homomorphic encryption scheme based on integer is proposed by Dijk, Gentry and Halevi, whose security is based on intractability of the approximate GCD and spare subset sum. A recent article [4] about fully homomorphic mainly improve the first fully homomorphic encryption scheme.

Obfuscation is a new research direction in the last few decades. Say in brief, obfuscation is a kind of algorithm, inputting a program, through the obfuscator, then output a new program which has the same function with the old one. Although the output circuit program on the function is same with the original program, but we find it difficult to read. Obfuscator like a "black box", in some degree, obfuscator will not leak any information on obfuscation program besides the input and output information.

The idea of this paper is to realize the obfuscation of homomorphic encryption through a point function obfuscator. According to one of the way of constructing point function obfuscation—perfect one-way hash function, the essence of perfect one-way hash function in [5] and [6] is point function obfuscator. The way of construction is based on hash function and computational difficulty problems. And now homomorphic encryption schemes are constructed based on computational difficulty problem. As both depend on computational difficulty problem. Then we'll learn the security obfuscation of homomorphic encryption according to the point function obfuscator constructed below.

In the rest of this paper, Sect. 2 presents the preliminaries of point function obfuscator; the point function obfuscator is constructed in Sect. 3; secure obfuscation scheme of homomorphic encryption is realized by this obfuscator is introduced in Sect. 4; finally, we prove the scheme's obfuscation character and security analysis in Sect. 5.

2 Preliminaries

2.1 The Introduction of Virtual Black-Box Obfuscation [7]

For circuit C, a probabilistic polynomial time O is a virtual black-box obfuscator. It meets three conditions as follows:

- (Functional) For any $n \in N, C \in C_n, O(C)$ is a circuit, and has the same computing function with C.
- (Polynomial slowdown) There exists a polynomial q, that for any $n \in N, C \in C_n, |O(C)| \leq q(|C|)$.
- (virtual black-box) For any PPT opponent A and polynomial p, there is a PPT simulator S. Then for all large enough $n \in N$ and $C \in C_n$:

$$\left| \underset{A,O}{pr}\left[A(O(C)) = 1\right] - \underset{S}{pr}\left[S^C\left(1^C\right) = 1\right] \right| \le {}^1\!/\mathrm{p(n)}$$

The obfuscator is effective if the obfuscator runs in polynomial time.

The black-box obfuscation in polynomial slowdown showing the complexity of the circuit and the virtual black-box performance provide a strong protection for the circuit security after obfuscated. In a sense, and it don't divulge any information about the circuit except its input and output behavior. More precisely, any valid adversary who can access to the obfuscation circuit can only get through an effective simulator which enters into the scheme through an oracle channel.

2.2 The Introduction of Semantic Perfect One-Way

For any non-unified PPT A and polynomial p, an ensemble $H = \{H^n\}_{n \in N}$, it is called semantic perfect one-way if it meets three conditions as follows:

Completeness: $\forall k \in K_n, x \in \{0, 1\}^n, r \in R_n, V(x, H_k(x, r)) = 1$
Collision resistance: For any non-unified PPT A:

$$\Pr[k \leftarrow K_n, (x_1, x_2, y) \leftarrow A(k) : x_1 \ne x_2 \wedge V(x_1, y) = V(x_2, y) = 1] < u(n)$$

Secrecy: There exists a non-unified PPTS, for sufficiently large n, any k and x:

$$|\Pr[r \leftarrow R_n, b \leftarrow A(k, H_k(x, r)) : b = 1]$$
$$- \Pr[r \leftarrow R_n, b \leftarrow S^{F_x}(k) : b = 1]| \le 1/p(n)$$

where F_x is a point function in x.

Note that semantic perfect one-way adopt to a simple way to virtual black box, the performance requirements of the obfuscation point function in the definition [8]. Therefore, a function which meets this definition is a point function obfuscation (possess approximate functional computing). But it is not real in the opposite direction. When semantic perfect one-way implies the virtual black box performance, the completeness and collision resistance in H will imply the approximate function of computing.

Definition 1 (*Homomorphic encryption*) [9] the encryption scheme $E = (KeyGen_\varepsilon, Encrypt_\varepsilon, Decrypt_\varepsilon, Evaluate_\varepsilon)$ is called homomorphism encryption scheme, if for every function f of a kind of specific function F_ε. The output ciphertext of $Evaluate_\varepsilon$ meet correctness requirements. Let such set of function F_ε equal $c_\varepsilon(\lambda)$.

Correctness: For any given security parameter λ, $KeyGen_\varepsilon(\lambda)$ output any key pairs $KeyGen_\varepsilon(\lambda)$, any $f \in F_\varepsilon$, any given plaintext $m_1, m_2, \ldots, m_t$ and the corresponding ciphertext $\vec{c} = (c_1, c_2, \ldots, c_t)$, and $c_i \leftarrow Encrypt_\varepsilon(pk, m_i)$, as $c \leftarrow Evaluat_\varepsilon(pk, f, \vec{c})$, then $Decrypt_\varepsilon(sk, c) = f(m_1, \ldots, m_t)$ established.

3 Construct k, rt^k Obfuscators

The construction process is as follows. Supposing p is a large safe prime, that is to say, $p = aq + 1$, where a is a little integer (For simplicity, we assume that $a = 2$). In Z_p^*, supposing Q is a subset of q time. (In other words, Q is a group of square model p). After input m and secret random input $k \in_R Q$, oracle hash function H computes $r = h(m)$ first, here h is a collision resistance hash function. Then output $H(m, k) = k$, rt^k (The calculation result is model p). Authentication algorithm V is simple; given an input m and a hash value $\langle a, b \rangle$, calculate $x = h(m)$ and accept it if $rt^a = b$.

The further description of this obfuscator is as follows:

(Construct k, rt^k point obfuscators) Suppose $g = \{G\}$ is a group overall, each G_n is the group whose prime order is p. We define an obfuscator O, for point in domain Z_p^*, there is under type: $C_x \longrightarrow Oc(k, rt^k)$, $k \xleftarrow{U} G^*$ is a random generator of G_n and $c(k, rt^k)$ is a circuit which input r, check whether $xt^a = rt^a$.

Under the strong variant of the decided Diffie-Hellman assumption, this construct is safe. This construction and the point obfuscator in [10] is semantic security, which based on logarithm of intractable problems over finite fields.

The introduction of point function obfuscation was proposed in [11] for the first time. More detailed introduction of point function obfuscation and two types of point function obfuscator constructed are given in [12] and [13]. The obfuscators above are constructed reference to the perfect one-way hash function.

We analyze this construction based on the strong Diffie-Hellman assumption variant which is used to reveal, and this construction meet the oracle security of random input and prior message.

Assumption 1 The Diffie-Hellman Indistinguishability Assumptions: Let k be a security parameter. Let $p = 2q + 1$ be a randomly chosen k-bit safe prime and let $g \in_R Q$ (where Q is the group of squares modulus p).

DHI Assumption I: Let $a, b, c \in_R Z_q^*$, then $\langle g^a, g^b, g^{ab} \rangle \approx \langle g^a, g^b, g^c \rangle$. DHI Assumption II: For any well-spread distribution ensemble $\{X_q\}$ where the domain of X_q is Z_q^*, for a drawn from X_q and for $b, c \in_R Z_q^*$ we have $\langle g^a, g^b, g^{ab} \rangle \approx \langle g^a, g^b, g^c \rangle$.

DHI Assumption III: For any uninvertible function f and for $a, b, c \in_R Z_q^*$ we have $\langle f(a), g^b, g^{ab} \rangle \approx \langle f(a), g^b, g^c \rangle$.

1. It can be seen that Assumption III implies Assumption II, and Assumption II implies Assumption I. We were unable to show implications in the other direction.
2. While these assumptions are considerably stronger than the standard Diffie-Hellman assumption (For p, g, g^a, g^b, it is only assumed that g^{ab} cannot be computed), they seem consistent with the current knowledge on the Diffie-Hellman problem. In particular, the assumption in the past is explicitly and

implicitly. It is not hard to see that it is equivalent to the semantic security of the ELGamal encryption scheme, because both of them are based on logarithm of intractable problems over finite fields.

Although Assumption II and III look quite strong, we were unable to contradict them. We propose the viability of these assumptions as an open question. To gain assurance in the plausibility of these assumptions, we remark that it is a common practice to use Diffie-Hellman key exchange modulo a large prime.

Theorem 1 *If DHI Assumption* (I, II, III) *holds then the function* $H(m, k) = k, rt^k$, *together with its verification algorithm, are an oracle hashing scheme for random inputs.*

4 Achieve a Security Obfuscator of Homomorphic Scheme by Obfuscator

A homomorphic encryption scheme is showed at first in this part (scheme 1), next we obfuscate this scheme (scheme 2). The two schemes is equal in the function, comparatively speaking, the readability of scheme 2 is worse (that is to say, "Hard to identify").

Scheme 1: A simple homomorphic encryption scheme (multiplication homomorphism)

1. Select a primitive element g, $g \in Z_p^*$;
2. Select an integer d randomly, $1 \le d \le p - 2$, compute $t = g^d \pmod{p}$;
3. Encryption transformation: For any plaintext $m \in Z_p^*$, secretly choose an integer k, $k \in_R Q$, then compute $E(m) = (g^k, mt^k)$;
4. Evaluate transformation: After obfuscation encrypt for any plaintext messages m_1, m_2,

$$E(m_1) \cdot E(m_2) (g^k, m_1 t^k)(g^k, m_2 t^k) = (g^{k_1+k_2}, ((m_1 \cdot m_2)t^{k_1+k_2})) = E(m_1 \cdot m_2)$$

5. Decryption transformation: Let

$$c_1 = g^{k_1+k_2} \bmod p, \ c_2 = (m_1 \cdot m_2)t^{k_1+k_2}, \ t = g^d \pmod{p},$$

$$c_2 (c_1^d)^{-1} \equiv (m_1 \cdot m_2)t^{k_1+k_2} \left(\left(g^{k_1+k_2} \right)^d \right)^{-1}$$

$$\equiv (m_1 \cdot m_2) g^{d(k_1+k_2)} \left(g^{-d(k_1+k_2)} \right) \pmod{p} \equiv (m_1 \cdot m_2)$$

Notice: p, g are the public elements, t is the encryption key and d is the confidential decryption key.

Scheme 2: We achieve security obfuscation of scheme 1 by the obfuscator which constructed in part 3 and retain the homomorphic characteristic of the scheme in the same time. The scheme as follows:

1. Select a primitive element g, $g \in Z_p^*$;
2. Select an integer d randomly, $1 \leq d \leq p - 2$, compute $t = g^d \pmod{p}$;
3. Encryption transformation: For any plaintext $m \in Z_p^*$, secret choose an integer k, $k \in_R Q$, compute $r = h(m)$, $E(r) = (g^k, rt^k)$;
4. Evaluate transformation: After obfuscation encrypt for any plaintext message m_1, m_2,

$$E(r_1) \cdot E(r_2) = (g^k, h(m_1)t^k)(g^k, h(m_2)t^k)$$
$$= (g^{k_1+k_2}, (h(m_1) \cdot h(m_2)t^{k_1+k_2})) = E(r_1 \cdot r_2)$$

5. Decryption transformation: Let $c_1 = g^{k_1+k_2} \bmod p$,

$$c_2 = (h(m_1) \cdot h(m_2))t^{k_1+k_2}, \, t = g^d \pmod{p}$$
$$c_2(c_1^d)^{-1} \equiv (h(m_1) \cdot h(m_2))t^{k_1+k_2}\left(\left(g^{k_1+k_2}\right)^d\right)^{-1}$$
$$\equiv (h(m_1) \cdot h(m_2))g^{d(k_1+k_2)}\left(g^{-d(k_1+k_2)}\right)\pmod{p} \equiv (h(m_1) \cdot h(m_2))$$

Notice: p, g are the public elements, t is the encryption key and d is the confidential decryption key.

According to the two schemes above, different ciphertext multiply after they encrypt different messages, which united equal to encryption of a set of different message. Then we find the set of message can be decryption by the decryption key. So the correctness of homomorphic encryption scheme is verified.

5 Proof and Analyze the Security of Secure Obfuscation of Homomorphic Encryption Scheme

In this part, we verify the correctness of homomorphic, then we prove the homomorphic encryption scheme of obfuscation security above is obfuscation secure and the mainly method is reducibility certification.

5.1 The Correctness of (Mainly for Homomorphic Encryption Scheme) Homomorphic

1. There exists a pair of key (t, d), t is the encryption key and d is the confidential decryption key.

2. After obfuscation encrypt for any plaintext message m_1, m_2, $c_{11} = g^k \mod p$, $c_{21} = (m_1)t^{k_1}$ and $c_{12} = g^k \mod p$, $c_{22} = (m_2)t^{k_2}$

3. According to evaluate transformation $c = \left(g^{k_1+k_2}, \left((m_1 \cdot m_2)t^{k_1+k_2}\right)\right)$;

4. $c_2\left(c_1^d\right)^{-1} \equiv (m_1 \cdot m_2)g^{d(k_1+k_2)}\left(g^{-d(k_1+k_2)}\right)(\bmod\, p) \equiv (m_1 \cdot m_2)$

The correctness verification of scheme 2 is the same as above, it can be set up according to the correctness verification of scheme 1, so the correctness of scheme 2 can be set up, i.e, the homomorphism encryption scheme after obfuscation will not change homomorphism.

5.2 Obfuscation Proof of the Scheme

Through the reducibility certification, it actually can be summed up in obfuscation security certification of the point function obfuscator in part 3. Now we divide it into two steps: the first, certificating this obfuscator meet three characteristics of the definition of obfuscation; the second, certificating this construction is perfect one-way hash function of the probability.

Proof The following will prove that this function satisfies the definition of obfuscator. (1) As we compute $r = h(m)$ at first, for a circuit C (the characteristic is mt^a), the result of obfuscation is $O(C)$, the two on the function of the computing can not be distinguished (keeping function). (2) It is cleared that the polynomial slowdown performance is established. (3) For any adversary A of probability polynomial time, there are a probability polynomial time S and a negligible function α, so for all the circuits C: $\left|\Pr[A(O(f)) = 1] - \Pr\left[S^C\left(1^{|C|}\right)\right]\right| \leq \alpha(|C|)$ (Characteristics of weak virtual black box).

This obfuscator is efficient if this obfuscator O runs in polynomial time. After the hash function processes the message, it reaches the difficulty identify characteristics of obfuscation, thus this construction is established to meet the obfuscator.

We will prove that this obfuscator is semantic perfect one-way. An ensemble $H = \{H^n\}_{n\in N}$, there is a decided polynomial time algorithm V. There exist a private key β in the key space, $x \in \{0, 1\}^n$, k is belong to a random element.

Because $h(x)\beta^k = H_\beta(x, k)$, so $V\left(x, H_\beta(x, k)\right) = 1$. (Completeness established). For any nonuniform probability polynomial time A and n which is large enough, there exists a negligible function $u(n)$ and message $x_1 \neq x_2$:$\Pr[\beta \in K_n, (x_1, x_2, y)$ $\leftarrow A(\beta) : V(x_1, y) = V(x_2, y) = 1] < u(n)$ (Collision resist-ance established).

For any nonuniform probability polynomial time A and polynomial p, there is a nonuniform probability polynomial time simulator S, for large enough n, any β and message x, input a $k \in R_n$ randomly:

$$\left| \Pr\left[b \leftarrow A\left(\beta, H_\beta(x,k) \right) : b = 1 \right] - \Pr\left[b \leftarrow S^{F_x}(\beta) : b = 1 \right] \right| \leq {}^1/_{p(n)}$$

Here F_x is a point function of independent variable x, therefore, the construction is semantic perfect one-wayness.

In more detail, let H be a semantic POW function. To obfuscate F_x, sample a seed, k, and random string, r, for H and output the obfuscation, $O(F_x) = k, H_k(x,r)$. The new function, $O(F_x)$, simply computes the predicate $V(., H_k(x,r))$. It can be shown that O is an obfuscator for the class of point functions. Completeness and collision resistance on H imply computational approximate functionality while semantic perfect one-wayness implies the virtual-black box property. On the other hand, an obfuscation of point functions may not be a POW function because approximate functionality does not rule out collisions chosen in an adversarial way.

In the definition of obfuscation, for the obfuscation of point function, we take a simple way and find that the semantic perfect one-way accord with the performance of virtual black box. Hence it is a obfuscation of point function if this function meets the semantic perfect one-wayness (The approximate function of computing). But it is not established in the opposite side.

6 Conclusion

In order to improve security, this paper in view of the existing homomorphic encryption scheme and propose homomorphic scheme of obfuscation security. This scheme is safer than other scheme and is difficult to identify. This article is the first time adds obfuscation characteristic to homomorphic public-key encryption and it will greatly protect the encryption data of user in cloud. Also this scheme can use in electronic ballot, which can improve the fairness of election. For the scheme, there are other purposes to be discovered.

References

1. Rivest RL, Adleman L, Dertouzos ML (1978) On data banks and privacy homeomorphi-sms. Found Secure Comput, pp 169–177
2. Gentry C (2009) Fully homomorphic encryption using ideal lattices. ACM STOC, pp 169–178
3. Gentry C (2009) A fully homomorphic encryption scheme. Ph.D. Thesis, Stanford University
4. Dijk MV, Gentry C, Halevi S, Vaikuntanathan V (2010) Fully homomorphic encryption over the integers. EUROCRYPT 2010, LNCS 6110, pp 24–43
5. Canetti R (1997) Towards realizing random oracles: hash functions that hide all partial information. In: Kaliski Jr BS (ed) CRYPTO 1997. LNCS, vol 1294. Springer, Heidelberg, pp 455–469

6. Canetti R, Micciancio D, Reingold O (1998) Perfectly one-way probabilistic hash functions. In: 30th ACM symposium on theory of computing, ACM. Press, pp 131–140

7. Barak B, Goldreich O, Impagliazzo R, Rudich S, Sahai A, Vadhan SP,Yang K (2001) On the (im)possibility of obfuscating program. CRYPTO 2001. pp 1–18

8. Canetti R, Dakdouk RR (2008) Obfuscating point functions with multibit output. In: Smart NP (ed) EUROCRYPT 2008. LNCS, vol. 4965, pp. 489–508. Springe-r,Heidelberg (2008)

9. Zhou Y (2010) Homomrphic cryptography research progress. China cryptography development report 2010. Publishing House of Electronics Industry, Beijing, pp 169

10. Bitansky N, Canetti R (2010) On strong simulation and composable point obfuscation. In: CRYPTO, advances in cryptology-CRYPTO 2010, 30th annual cryptology conference, Santa Barbara, CA, USA, 15–19 August, 2010, Proceedings, pp 520–537

11. Lynn BYS, Prabhakaran M, Sahai A (2004) Positive results and techniques for obfuscation. In: Cachin C, Camenisch JL (eds) EUROCRYPT 2004. LNCS, vol 3027. Springer, Heidelberg, pp 20–39

12. Wee H (2005) On obfuscation point functions. In: Proceedings of the 37th ACM symposium on theory of computing, pp 523–532

13. Canetti R, Kalai YT, Varia M, Wichs D (2010) On symmetric encryption and point obfuscation. In: Micciancio D (ed) TCC 2010. LNCS, vol 5978. Springer, Heidelberg, pp 52–71

Construction of Barycentric Blending Rational Interpolation Over the Triangular Grids

Qiang Li and Feng Xu

Abstract Laying down the foundation for the basic function of the barycentric rational interpolation, some rational interpolations over all kinds of triangle grids were constructed, and duality theorems and characterization theorems were given, some significative characters are obtained. Compared with the traditional rational interpolation based on continued fraction, the barycentric blending interpolation inherited the advantages of the simple expressions, has many advantages such as small calculation quantity, good numerical stability, no poles and unattainable points, etc. The barycentric blending interpolation can also be extended to both higher dimensions, vector-valued case and matrix-valued case.

Keywords Barycentric rational interpolation · Duality theorems · Characterization theorems

1 Introduction

Polynomial interpolants are used as the basic means of approximation in nearly all areas of numerical analysis. Examples of meromorphic functions for which the polynomial interpolant does not converge uniformly were given by Meray and later Runge, which limited application of the interpolation polynomial. It is well known that the classical rational interpolation sometimes gives better approximations than polynomial interpolation, especially for large sequences of points. But it is difficult to avoid and control poles and there is sometimes unattainable

Q. Li (✉) · F. Xu
College of Science, Anhui University of Science and Technology, Huainan 232001, China
e-mail: liqiangtw@sina.com

F. Xu
e-mail: xufeng19631223@126.com

points and infinite inverse differences for Thiele-type continued fraction interpolation [1, 2].

Barycentric rational interpolation was presented by W. Werner, which possess various advantages in comparison with classical continued fraction rational interpolants, such as barycentric rational interpolants have small calculated amount, good numerical stability, poles and unattainable points are prevented through choose weights, regardless of the distribution of the points [3–6]. In this paper, based on Thiele-type continued fraction interpolants and Barycentric rational interpolants, new bivariste blending rational interpolation are over the triangular grids are constructed, and duality theorems and characterization theorems were given [7, 8].

Suppose $\prod^{n,n}$ is a given square set of points which is arranged in the following array:

$$
\begin{array}{ccc}
(x_0, y_0) & \cdots & (x_0, y_n) \\
\vdots & \ddots & \vdots \\
(x_n, y_0) & \cdots & (x_n, y_n)
\end{array}
\tag{1}
$$

where $x_i \neq x_j$, $y_i \neq y_j$, if $i \neq j$. The square array (1) may be decomposed into the following four triangular sub-arrays (A), (B), (C) and (D) denoted by LB, RU, RB, LU, respectively.

$$
\begin{array}{cccc}
(x_0, y_0) & & & \\
(x_1, y_0) & (x_1, y_1) & & \\
\vdots & \vdots & \ddots & \\
(x_n, y_0) & (x_n, y_1) & \cdots & (x_n, y_n) \\
& & (A) &
\end{array}
\qquad
\begin{array}{cccc}
(x_0, y_0) & (x_0, y_1) & \cdots & (x_0, y_n) \\
& (x_1, y_1) & \cdots & (x_1, y_n) \\
& & \ddots & \vdots \\
& & & (x_n, y_n) \\
& (B) & &
\end{array}
$$

$$
\begin{array}{cccc}
& & (x_0, y_n) & \\
& (x_1, y_{n-1}) & (x_1, y_n) & \\
\cdot\cdot\cdot & \vdots & \vdots & \\
(x_n, y_0) & \cdots & (x_n, y_{n-1}) & (x_n, y_n) \\
& & (C) &
\end{array}
\qquad
\begin{array}{ccc}
(x_0, y_0) & (x_0, y_1) & \cdots & (x_0, y_n) \\
\vdots & \vdots & \cdot\cdot\cdot \\
(x_{n-1}, y_0) & (x_{n-1}, y_1) & \\
(x_n, y_0) & & \\
& (D) &
\end{array}
$$

And (A), (B), (C), (D) denoted the left lower triangular grid (LB), the right upper triangular grid (RU), the left lower triangular grid (RB) and the left upper triangular grid (LU), respectively, namely,

$$
\begin{aligned}
LB &= \{(x_i, y_j) \,|\, j = 0, 1, \cdots i; \ i = 0, 1, \cdots n\}; \\
RU &= \{(x_i, y_j) \,|\, j = i+1, \cdots n; \ i = 0, 1, \cdots n\}; \\
RB &= \{(x_i, y_j) \,|\, j = n-i, \cdots, n; \ i = 0, 1, \cdots n\}; \\
LU &= \{(x_i, y_j) \,|\, j = 0, 1, \cdots n-i; \ i = 0, 1, \cdots n\}
\end{aligned}
$$

The rational interpolation problem consists in finding a rational function $R(x, y) = \frac{P(x,y)}{Q(x,y)}$ satisfying

$$R(x_i, y_j) = \frac{P(x_i, y_j)}{Q(x_i, y_j)} = f(x_i, y_j), \ (x_i, y_j) \in LB, RU, RB \text{ or } LU.$$

2 Construction of Barycentric Blending Rational Interpolation Over the Triangular Grids

We incorporate barycentric interpolation and Thiele-type continued fraction interpolants, to construct barycentric blending rational interpolation over the triangular grids

$$R(LB, x, y) = \frac{\sum_{i=0}^{m} \omega_i A_i(y) \prod_{k=0, k \neq i}^{m} (x - x_k)}{\sum_{i=0}^{m} \omega_i \prod_{k=0, k \neq i}^{m} (x - x_k)}, \ j = 0, 1, \cdots, i; \ i = 0, 1, \cdots, n \quad (2)$$

then, ω_i are the weights,.

$$A_i(y) = a_{i0} + \frac{y - y_0}{a_{i1}} + \frac{y - y_1}{a_{i2}} + \cdots + \frac{y - y_{n-1}}{a_{in}}, i = 0, \cdots, n$$

Definition 1 Let

$$\rho[x_i, y_j] = f_{ij}; \quad (3)$$

$$\rho[x_i, y_j, y_r] = \frac{y_r - y_j}{\rho[x_i; y_r] - \rho[x_i; y_j]}; \quad (4)$$

$$\rho[x_i, y_r, \cdots, y_s, y_k, y_l] = \frac{y_r - y_j}{\rho[x_i, y_r, \cdots, y_s, y_l] - \rho[x_i, y_r, \cdots, y_s, y_k]} \quad (5)$$

We call $a_{ij} = \rho[x_i, y_0, y_1, \cdots, y_j]$ define by formulas (3)–(5) as the inverse differences of $f(x, y)$ at the set of points LB.

Theorem 2.1 *Let* $\omega_i \neq 0$ *and* $a_{ij} = \rho[x_i, y_0, y_1, \cdots, y_j]$ *define by formulas* (3)–(5) *Then*

$$R(LB, x, y) = \frac{\sum_{i=0}^{m} \omega_i A_i(y) \prod_{k=0, k \neq i}^{m} (x - x_k)}{\sum_{i=0}^{m} \omega_i \prod_{k=0, k \neq i}^{m} (x - x_k)}, j = 0, 1, \cdots, i; \ i = 0, 1, \cdots, n$$

is an interpolation formula, have the following interpolation property

$$R\big(LB; x_i, y_j\big) = f\big(LB; x_i, y_j\big)(j = 0, 1, \cdots, n; \, i = j, j+1, \cdots, n).$$

Proof From (3) to (5), $\forall (x_i, y_j) \in \prod^{mn}$, it follows:

$$
\begin{aligned}
A_i(y_j) &= a_{i0} + \frac{y_j - y_0}{a_{i1}} + \frac{y_j - y_1}{a_{i2}} + \cdots + \frac{y_j - y_{j-1}}{a_{ij}} \\
&= \rho[x_i; y_0] + \frac{y_j - y_0}{\rho[x_i; y_0, y_1]} + \frac{y_j - y_1}{\rho[x_i; y_0, y_1, y_2]} + \cdots + \frac{y_j - y_{j-1}}{\rho\big[x_i; y_0, y_1, \cdots, y_j\big]} \\
&= \rho[x_i; y_0] + \frac{y_j - y_0}{\rho[x_i; y_0, y_1]} + \frac{y_j - y_1}{\rho[x_i; y_0, y_1, y_2]} + \cdots + \frac{y_j - y_{j-2}}{\rho\big[x_i; y_0, y_1, \cdots y_{j-2}, y_j\big]} \\
&= \rho[x_i; y_0] + \frac{y_j - y_0}{\rho\big[x_i; y_0, y_j\big]} \\
&= \rho\big[x_i; y_j\big] = f_{ij}, \; j = 0, 1, \cdots n; \, i = j, j+1, \cdots, n
\end{aligned}
$$

which leads to

$$
\begin{aligned}
R(x_i, y_j) &= \frac{\sum\limits_{t=0}^{m} \omega_t \prod_{k \neq 0, k \neq i}^{m} \big(x_i - x_j\big) f_{ij}}{\sum\limits_{t=0}^{m} \omega_t \prod_{k \neq 0, k \neq i}^{m} \big(x_i - x_j\big)} \\
&= \frac{\omega_i A_i\big(y_j\big)(x_i - x_0)(x_i - x_1) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_m)}{\omega_i (x_i - x_0)(x_i - x_1) \cdots (x_i - x_{i-1})(x_i - x_{i+1}) \cdots (x_i - x_m)} \\
&= f_{ij}, \; j = 0, 1, \cdots n; \, i = j, j+1, \cdots, n.
\end{aligned}
$$

as asserted. Similarly, one can prove the following theorems:

Theorem 2.2 *Let* $\omega_i \neq 0$ *and* $a_{ij} = \rho[x_i, y_0, y_1, \cdots, y_j]$ *define by formulas* (3)–(5). *Then*

$$R(RU, x, y) = \frac{\sum\limits_{i=0}^{m} \omega_i A_i(y) \prod\limits_{k=0, k \neq i}^{m} (x - x_k)}{\sum\limits_{i=0}^{m} \omega_i \prod\limits_{k=0, k \neq i}^{m} (x - x_k)}, \, j = i, i+1, \cdots, n; i = 0, 1, \cdots n$$

is an interpolation formula, have the following interpolation property.

Theorem 2.3 *Let* $\omega_i \neq 0$ *and* $a_{ij} = \rho[x_i, y_0, \cdots, y_j]$ *define by formulas* (3)–(5). *Then*

$$R(RB, x, y) = \frac{\sum\limits_{i=0}^{m} \omega_i A_i(y) \prod\limits_{k=0, k \neq i}^{m} (x - x_k)}{\sum\limits_{i=0}^{m} \omega_i \prod\limits_{k=0, k \neq i}^{m} (x - x_k)}, \, j = n - i, \cdots, n; i = 0, 1, \cdots n$$

is an interpolation formula, have the following interpolation property.

Theorem 2.4 *Let $\omega_i \neq 0$ and $a_{ij} = \rho[x_i, y_0, y_1, \cdots, y_j]$ define by formulas (3)–(5). Then*

$$R(LU, x, y) = \frac{\sum_{i=0}^{m} \omega_i A_i(y) \prod_{k=0,k\neq i}^{m} (x - x_k)}{\sum_{i=0}^{m} \omega_i \prod_{k=0,k\neq i}^{m} (x - x_k)}, j = 0, 1, \cdots n - i; i = 0, 1, \cdots n$$

is an interpolation formula, have the following interpolation property.

3 The Duality of Barycentric Blending Rational Interpolation Over the Triangular Grids

In this paper, the method and some conclusions in [8] are extend to study theduality interpolation formula

$$DR(LB, x, y) = \frac{\sum_{i=0}^{m} \omega_j A_j(x) \prod_{p=0,p\neq j}^{n} (y - x_p)}{\sum_{j=0}^{n} \omega_j \prod_{j=0,p\neq j}^{n} (y - x_p)}, j = 0, \cdots, i; i = 0, \cdots, n \quad (6)$$

Then, ω_j is weights,

$$B_i(x) = b_{0j} + \frac{x - x_0}{b_{1j}} + \frac{x - x_1}{a_{2j}} + \cdots + \frac{x - x_{m-1}}{a_{mj}}, j = 0, \cdots, n.$$

Definition 2 Let

$$\rho[x_i, y_j] = f_{ij}; \quad (7)$$

$$\rho[x_i, x_r, y_j] = \frac{x_r - x_i}{\rho[x_r; y_j] - \rho[x_i; y_j]}; \quad (8)$$

$$\rho[x_r, \cdots, x_s, x_k, x_l; y_j] = \frac{x_l - x_k}{\rho[x_r, \cdots, x_s, x_l; y_j] - \rho[x_r, \cdots, x_s, x_k; y_j]} \quad (9)$$

We call $b_{ij} = \rho[x_0, x_1, \cdots, x_i, y_j]$ define by formulas (7)–(9) as the inverse differences of $f(x, y)$ at the set of points LB. From formula (2) and (6), it is easy to derive the following duality theorem.

Theorem 3.1

$$R(LB; x_i, y_j) = DR(LB; x_i, y_j)(j = 0, \cdots, n; i = j, \cdots, n).$$

Proof According to Theorem 2.1, we may write

$$R\big(LB;x_i,y_j\big)=f\big(LB;x_i,y_j\big)(j=0,1,\cdots,n;\ i=j,j+1,\cdots,n),$$

from (7)–(9), $\forall\big(x_i,y_j\big)\in\prod^{mn}$, it follows

$$
\begin{aligned}
B_j(x_i)&=b_{0j}+\frac{x_i-x_0}{b_{1j}}+\frac{x_i-x_1}{b_{2j}}+\cdots+\frac{x_i-y_{i-1}}{b_{ij}}\\[4pt]
&=\rho\big[x_0;y_j\big]+\frac{x_i-x_0}{\rho\big[x_0;x_1,y_j\big]}+\frac{x_i-x_1}{\rho\big[x_0;x_1,x_2,y_j\big]}+\cdots+\frac{x_i-x_{i-1}}{\rho\big[x_0;x_1,x_2,\cdots,x_s,y_j\big]}\\[4pt]
&=\rho\big[x_i;y_j\big]+\frac{x_s-x_0}{\rho\big[x_0,x_1;y_j\big]}+\frac{x_s-x_1}{\rho\big[x_0,x_1,x_2;y_j\big]}+\cdots+\frac{x_s-x_{s-2}}{\rho\big[x_0,x_1,x_2,\cdots x_{s-2},x_s;y_j\big]}\\[4pt]
&=\rho\big[x_0;y_j\big]+\frac{x_s-x_0}{\rho\big[x_0;x_s,y_j\big]}\\[4pt]
&=\rho\big[x_i;y_j\big]=f_{ij},\ j=0,1,\cdots n;\ i=j,j+1,\cdots,n
\end{aligned}
$$

which leads to

$$
\begin{aligned}
DR(LB,x_i,y_j)&=\frac{\displaystyle\sum_{i=0}^{m}\omega_j A_j(x)\prod_{p=0,p\neq j}^{n}\big(y-x_p\big)}{\displaystyle\sum_{j=0}^{n}\omega_j\prod_{j=0,p\neq j}^{n}\big(y-x_p\big)}\\[8pt]
&=\frac{\omega_j b_j(x_i)\big(y_j-y_0\big)\big(y_j-y_1\big)\cdots\big(y_j-y_{j-1}\big)\big(y_j-y_{j+1}\big)\cdots\big(y_j-y_n\big)}{\omega_j\big(y_j-y_0\big)\big(y_j-y_1\big)\cdots\big(y_j-y_{j-1}\big)\big(y_j-y_{j+1}\big)\cdots\big(y_j-y_n\big)}\\[6pt]
&=f_{ij},\ j=0,1,\cdots n;\ i=j,j+1,\cdots,n
\end{aligned}
$$

then $R_{m,n}\big(LB,x_i,y_j\big)=DR_{m,n}\big(LB,x_i,y_j\big)=f\big(LB,x_i,y_j\big)$ as asserted.

Similarly, one can prove the following theorems:

Theorem 3.2 $R\big(RU;x_i,y_j\big)=DR\big(RU;x_i,y_j\big)(j=i,\cdots,n;i=0,\cdots n).$

Theorem 3.3 $R\big(RB;x_i,y_j\big)=f\big(RB;x_i,y_j\big)(j=n-i,\cdots,n;i=0,\cdots n).$

Theorem 3.4 $R\big(LU;x_i,y_j\big)=DR\big(LU;x_i,y_j\big)(j=0,1,\cdots n-i;\ i=0,\cdots n).$

4 Characterization Theorem of Barycentric Blending Rational Interpolation Over the Triangular Grids

In this paper, we based on the Ref. [2] studied the $R(LB,x,y)$, $DR(LB,x,y)$, and give the following theorems:

Theorem 4.1 *The function $R(LB,x,y)$ is of type $[m/m]$ with respect to x and of type $\big[\frac{n+1}{2}\big/\frac{n}{2}\big]$ with respect to y; The dual function $DR(LB,x,y)$ is of type $\big[\frac{m+1}{2}\big/\frac{m}{2}\big]$ with respect to x and of type $[n/n]$ with respect to y.*

Proof According to Ref. [2], it is easy to know that the blending rational interpolant $A_i(y)$ is of type $\left[\frac{n+1}{2}\big/\frac{n}{2}\right]$ with respect to y, $\sum_{i=0}^{m} \prod_{k=0,k\neq i}^{m} (x-x_k)$ has degree m about x, and $\sum_{i=0}^{m} \omega_i \prod_{k=0,k\neq i}^{m} (x-x_k)$ has degree about x, we can draw the conclusion that the blending rational interpolant $R(LB,x,y)$ is of type $[m/m]$ with respect to x and of type $\left[\frac{n+1}{2}\big/\frac{n}{2}\right]$ with respect to y. While the dual function $DR(LB,x,y)$ is of type $\left[\frac{m+1}{2}\big/\frac{m}{2}\right]$ with respect to x and of type $[n/n]$ with respect to y as asserted.

Similarly, one can prove the following theorems:

Theorem 4.2 *The function $R(RU,x,y)$ is of type $[m/m]$ with respect to x and of type $\left[\frac{n+1}{2}\big/\frac{n}{2}\right]$ with respect to y; The dual function $DR(RU,x,y)$ is of type $\left[\frac{m+1}{2}\big/\frac{m}{2}\right]$ with respect to x and of type $[n/n]$ with respect to y.*

Theorem 4.3 *The function $R(RB,x,y)$ is of type $[m/m]$ with respect to x and of type $\left[\frac{n+1}{2}\big/\frac{n}{2}\right]$ with respect to y; The dual function $DR(RB,x,y)$ is of type $\left[\frac{m+1}{2}\big/\frac{m}{2}\right]$ with respect to x and of type $[n/n]$ with respect to y.*

Theorem 4.4 *The function $R(LU,x,y)$ is of type $[m/m]$ with respect to x and of type $\left[\frac{n+1}{2}\big/\frac{n}{2}\right]$ with respect to y; The dual function $DR(LU,x,y)$ is of type $\left[\frac{m+1}{2}\big/\frac{m}{2}\right]$ with respect to x and of type $[n/n]$ with respect to y.*

5 Conclusion

In this paper, based on Thiele-type continued fraction interpolants and Barycentric rational interpolants, new bivariste blending rational interpolation are over the triangular grids are constructed, and duality theorems and characterization theorems were given. The new rational interpolation inherited the advantages of the simple expressions, easy to calculate of continued fraction and small calculation quantity, no poles, Good numerical stability of barycentric rational interpolants. And the technique can also be extended to both higher dimensions, vector-valued case and matrix-valued case.

References

1. Renhong W (2007) The approximation of rational fuctions and applications. Science Press, Beijing (in Chinese)
2. Tan J (2007) Continued fraction theory and its application. Science Press, Beijing (in Chinese)

3. Schneider C, Werner W (1986) Some new aspects of rational interpolation. Math Compt 175(47):285–299
4. Floater MS, Hormann K (2007) Barycentric rational interpolation with no poles and high rates of approximations. Numberische Mathematik 107:315–331
5. Berrut JP, Mitemann HD (1997) Lebesgue constant minimizing linear rational interpolation of continuous functions over the interval. Comput Appl Math 33(6):77–86
6. Berrut JP, Trefethen LN (2004) Barycentric lagrange interpolation. SIAM Rev 46:501–517
7. Shen X (2011) Barycentric-Thiele type blending rational interpolants over rectangula grids. Sciencepaper online 6(10):726–731
8. Gong-qin Z (1995) The duality of bivariate vector valued rational interpolations over rectangular grids. Numer Math Sin 17(3):311–320

Hopf Bifurcation Analysis in an Intracellular Calcium Oscillation Model

Yuanhua Li, Zhou Yi and Hongkun Zuo

Abstract Besides the mechanism of the calcium-induced calcium release (CICR), the Kummer-Olsen Calcium oscillation model focuses on the effect of the feedback inhibition on the initial agonist receptor complex by calcium and activated phospholipase C, as well as receptor type-dependent self-enhanced behavior of the activated G_α subunit. The nonlinear dynamics of this model are investigated based on the theory of center manifold, bifurcations and stability of the equilibrium. Computations show that both appearance and disappearance of calcium oscillations of this system are related to supercritical Hopf bifurcation of the equilibrium. Numerical simulations including bifurcation diagram, temporal evolution and phase portraits, are performed to confirm the theoretical analysis. These results may be instructive for understanding the mechanism of other similar calcium oscillation models.

Keywords Hopf bifurcation · Calcium oscillation · Center manifold

1 Introduction

Many biological functions, from egg fertilization to cell death, are controlled by the oscillatory changing of free cytoplasmic calcium concentration [1, 2]. Calcium oscillations were discovered experimentally by Woods et al. in 1987, and large

Y. Li · Z. Yi · H. Zuo (✉)
Department of Math, Huainan Normal University, Huainan 232038, Anhui,
People's Republic of China
e-mail: navyhero@yahoo.cn

Y. Li
e-mail: hnliyuanhua@163.com

Z. Yi
e-mail: zhouyi3280@163.com

numbers of cells show calcium oscillations after simulated by an extracellular agonist. Lots of experimental works have confirmed the significant role of oscillations, and then many mathematical models were proposed in order to explain the mechanism of Ca^{2+} oscillations.

Shen and Larter [3] proposed a model of Ca^{2+} oscillations, which based on both the calcium-induced calcium release (CICR) and the inositol trisphosphate (IP_3) cross coupling (ICC), and numerical simulation was carried out to explain the mechanism of bursting oscillation. A more detailed study of the complex calcium oscillations in non-excitable cells was given by Borghans and Houart et al. [4–6]. Zhang and Chang et al. analyzed Li-Rinzel Ca^{2+} oscillation model by using the theories in dynamical systems and concluded that the oscillation is due to supercritical Hopf bifurcation and subcritical Hopf bifurcation.

In this article, we analyzed the Kummer-Olsen Ca^{2+} oscillation model by applying the stability theory and bifurcation theory and we were concerned with the types and stability of the equilibrium. By center manifold theorem and bifurcation theory [7–9], oscillations related to two supercritical Hopf bifurcations can be obtained. Finally, with the help of Matlab we perform some numerical simulations to confirm our conclusions.

2 Model Description

The model proposed by Kummer et al. focuses on the feedback inhibition on the initial agonist receptor complex by Ca^{2+} and activated phospholipase C (PLC) and receptor type-dependent self-enhanced behavior of the activated G_α subunit. So the three main variables are: the free Ca^{2+} concentration in the cytosol (Ca_{cyt}), the concentration of the active G_α subunit (G_α) and the concentration of active PLC (PLC). The model can be described by the following differential equations:

$$\begin{cases} dCa_{cyt}/dt = k_{10}G_\alpha - k_{11}Ca_{cyt}/(Ca_{cyt} + k_{12}) \\ dG_\alpha/dt = k_1 + k_2 G_\alpha - k_3 PLC G_\alpha/(G_\alpha + k_4) - k_5 Ca_{cyt} G_\alpha/(G_\alpha + k_6); \\ dPLC/dt = k_7 G_\alpha - k_8 PLC/(PLC + k_9) \end{cases} \quad (1)$$

other parameters are: $k_1 = 0.212$, $k_2 = 1.5$, $k_4 = 0.19$, $k_5 = 4.88$, $k_6 = 1.18$, $k_7 = 1.24$, $k_8 = 32$, $k_9 = 29$, $k_{10} = 13.58$, $k_{11} = 153$, $k_{12} = 0.16$ and the meaning of each parameter can be referred in [4]. In the following analysis k_3 is used as bifurcation parameter.

3 Analysis of Stability and Bifurcation of Equilibrium

Let $x = Ca_{cyt}$, $y = G_\alpha$, $z = PLC$, $r = k_3$ System (1) can be rewritten in the following form:

$$\begin{cases} dx/dt = 13.58y - 153x/(x + 0.16) \\ dy/dt = 1.5y - 4.88xy/(y + 1.18) - ryz/(y + 0.19) + 0.212 \\ dz/dt = 1.24y - 32z/(z + 29) \end{cases} \quad (2)$$

According to the actual meaning of x, y, z and r, only one condition needs to be studied whether there exists equilibrium of system (2) when $r \in [0.2, 2.4]$.

Based on theoretical analysis through the qualitative theory and stability theory we can obtain:

When $r < 0.788038$, the equilibrium is a stable node; when $r = 0.788038$, the system (2) has a non-hyperbolic equilibrium $O_1 = (0.4721939, 8.415148, 14.03226)$; when $0.788038 < r < 2.017361$, the system (2) has a saddle; when $r = 2.017361$, the system (2) has a non-hyperbolic equilibrium $O_2 = (0.009802158, 0.6503844, 0.7497654)$; when $r > 2.017361$, there exists a stable node of system (2).

From the above analysis, we can infer that the system (2) has a non-hyperbolic equilibrium O_1 (r_1) and O_2 (r_2), for $r_1 = 0.788038$ and $r_2 = 2.017361$ respectively. Now dynamics around these two equilibria are analyzed by using the center manifold theorem and bifurcation theory.

When $r = r_0$, the corresponding equilibrium of system (2) is (x_0, y_0, z_0), and let $x_1 = x - x_0$, $y_1 = y - y_0$, $z_1 = z - z_0$, $r_1 = r - r_0$. In order to apply the center manifold theorem with parameters of a continuous dynamical system, r_1 is regard as a new dynamical variable. Under the consideration of $dr_1/dt = 0$ and system (2), we can get:

$$\begin{cases} dx_1/dt = 13.58(y_1 + y_0) - 153(x_1 + x_0)/(x_1 + x_0 + 0.16) \\ dy_1/dt = 1.5(y_1 + y_0) - 4.88(x_1 + x_0)(y_1 + y_0)/(y_1 + y_0 + 1.18) \\ \qquad - (r_1 + r_0)(y_1 + y_0)(z_1 + z_0)/(y_1 + y_0 + 0.19) + 0.212 \, ; \\ dz_1/dt = 1.24(y_1 + y_0) - 32(z_1 + z_0)/(z_1 + z_0 + 29) \\ dr_1/dt = 0 \end{cases} \quad (3)$$

It's obvious that the origin $O(x_1, y_1, z_1, r_1) = (0, 0, 0, 0)$ is the equilibrium of system (3) when $r_1 = 0$, which has a same conclusion as the one (x_0, y_0, z_0) of system (2) through the analysis in type, stability and bifurcation type. The characteristic equation associated with system (3) is given by

$$P(\xi) = \xi(\xi^3 + Q_1\xi^2 + Q_2\xi + Q_3) = 0; \quad (4)$$

where $Q_1 = -(A + E + I)$, $Q_2 = AE + AI + EI - BD - FH - CG$, $Q_3 = BDE + FHA + CEG - AEI - BFG - DCH$, $A = -24.48/(x_0 + 0.16)^2$, $B = 13.58$, $C = 0$, $D = -4.88y_0/(y_0 + 1.18)$, $E = 1.5 - 0.19r_0z_0/(y_0 + 0.19)^2 - 5.7584x_0/(y_0 + 1.18)^2$, $F = -r_0 y_0/(y_0 + 0.19)$, $G = 0$, $H = 1.24$, $I = -928/(z_0 + 29)^2$.

When $r_0 = 0.788038$, characteristic roots of equilibrium point $O_2 = (0, 0, 0, 0)$ of system (3) are, $\xi_1 = -60.3097$, $\xi_2 = 0.8482i$, $\xi_3 = -0.8482i$, $\xi_4 = 0$, respectively.

$$\text{Let} \begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ r_1 \end{pmatrix} = U \begin{pmatrix} u \\ v \\ w \\ s \end{pmatrix}, \ U = \begin{pmatrix} -0.9976 & -0.0711 & 0.1166 & -0.0827 \\ -0.0691 & -0.3135 & 0.5306 & -0.3731 \\ 0.0014 & -0.7756 & 0 & -0.9233 \\ 0 & 0 & 0 & 0.0384 \end{pmatrix},$$

the system (3) becomes

$$\begin{pmatrix} \dot{u} \\ \dot{v} \\ \dot{w} \\ \dot{s} \end{pmatrix} = \begin{pmatrix} -60.3097 & 0 & 0 & 0 \\ 0 & 0 & -0.8482 & 0 \\ 0 & 0.8482 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \\ w \\ s \end{pmatrix} + \begin{pmatrix} g_1 \\ g_2 \\ g_3 \\ g_4 \end{pmatrix} \qquad (5)$$

where

$f_1 = 13.58g_{12} - 153g_{11}/(g_{11} + 0.16);$
$f_2 = 0.212 + 1.5g_{12} - 0.0384sg_{12}\,g_{13}/(g_{12} + 0.19) - 4.88g_{11}g_{12}/(g_{12} + 1.18);$
$f_3 = 1.24g_{12} - 32g_{13}/(g_{13} + 29);$
$g_1 = -1.0179f_1 + 0.2237f_2 + 0.0029f_3 + 60.3097u;$
$g_2 = -0.0018f_1 + 0.0004f_2 - 1.2893f_3 + 0.8482w;$
$g_3 = -0.1336f_1 + 1.914f_2 - 0.7614f_3 - 0.8482v;$
$g_4 = 0; \ g_{11} = -0.9976u - 0.0711v + 0.1166w - 0.0827s + 0.4721939;$
$g_{12} = -0.0691u - 0.3135v + 0.5306w - 0.3731s + 8.415148;$
$g_{13} = 0.0014u - 0.7756v - 0.9233s + 14.03226.$

$$(6)$$

By using the center manifold theory, we can infer that there exists a center manifold of system (5), and its form can be expressed as follows:

$$W_{loc}^c(O_1) = \{(u, v, w, s) \in R^4 \,|\, u = h^*(v, w, s), h^*(0, 0, 0) = 0, \ Dh^*(0, 0, 0) = 0\} \qquad (7)$$

let $h^*(v, w, s) = a_1v^2 + bw^2 + cs^2 + d_1vw + evs + fws + \cdots$, the center manifold of system (5) should meet

$$Dh^* \cdot (-0.8482w + g_2, 0.8482v + g_3, 0)^T + 60.3097\,h^* = g_1 \qquad (8)$$

From (8) we know that $a_1 = -0.0081$, $b = -0.0218$, $c = -0.0104$, $d_1 = 0.0267$, $e = -0.0184$, $f = 0.0303$. So the system confined by this center manifold of system (5) is

$$\begin{pmatrix} \dot{v} \\ \dot{w} \end{pmatrix} = \begin{pmatrix} 0 & -0.8482 \\ 0.8482 & 0 \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix} + \begin{pmatrix} f^1(v, w) \\ f^2(v, w) \end{pmatrix} \qquad (9)$$

where
$f^1(v, w) = -0.009953v^2 + 0.00302vw - 0.002476w^2 - 0.00082vw^2 + \cdots;$
$f^2(v, w) = -0.0729v^2 + 0.22169vw - 0.18184w^2 - 0.0384v^2w - 0.0344w^3 + \cdots$
Hence, it can be concluded that:

$$a = \frac{1}{16}[f_{vvv}^1 + f_{vww}^1 + f_{vvw}^2 + f_{wwww}^2]\big|_{(0,0)} + \frac{1}{16 \times 0.8482}[f_{vw}^1(f_{vv}^1 + f_{ww}^1)$$

$$- f_{vw}^2(f_{vv}^2 + f_{ww}^2) - f_{vv}^1 f_{vv}^2 + f_{ww}^1 f_{ww}^2)]\big|_{(0,0)} = -0.034047 < 0,$$

$$d = \frac{d(\text{Re}(\xi(s)))}{ds}\bigg|_{(0,0,0)} = 270.5136 > 0$$

To verify the analysis we need to summarize the discussion via the Hopf bifurcation theory.

Conclusion 1: A supercritical Hopf bifurcation occurs when r passes through $r_0 = 0.788038$ of system (2). When $r < r_0$, the equilibrium O_1 is stable, while $r > r_0$, the equilibrium O_1 will lose its stability. At the same time stable periodic solutions arise around the unstable equilibrium and the system (2) begins to oscillate.

When $r_0 = 2.017361$, the system confined by the center manifold of system (3) is

$$\begin{pmatrix} \dot{v} \\ \dot{w} \end{pmatrix} = \begin{pmatrix} 0 & -0.9147 \\ 0.9147 & 0 \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix} + \begin{pmatrix} f^1(v, w) \\ f^2(v, w) \end{pmatrix} \tag{10}$$

where
$$f^1(v, w) = -0.01552v^2 + 0.003977v^3 + \cdots;$$
$$f^2(v, w) = 0.03805v^2 + 0.03534vw + 0.01541w^2 - 0.00089v^2w$$
$$-0.000227w^3 + \cdots.$$

The following conclusions can be inferred when $a = -0.0009583 < 0$ and $d = -14282.4472 < 0$.

Conclusion 2: A supercritical Hopf bifurcation occurs when r passes through $r_0 = 2.017361$ of system (2). When $r < r_0$, the equilibrium O_2 of system (2) is unstable, and system (2) is going to oscillate. However, when $r > r_0$, the equilibrium point of system (2) is stable, and oscillations cease.

Fig. 1 Equilibria curve of system (2) in (r, y) plane

(a)

(b)
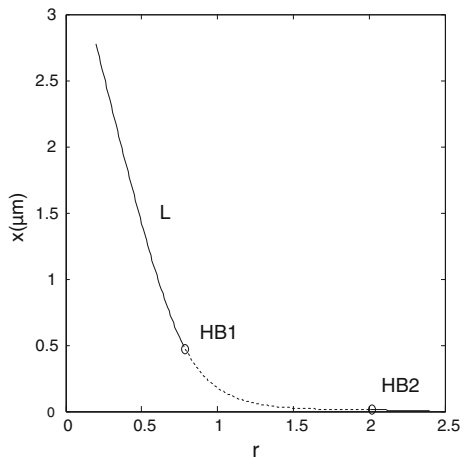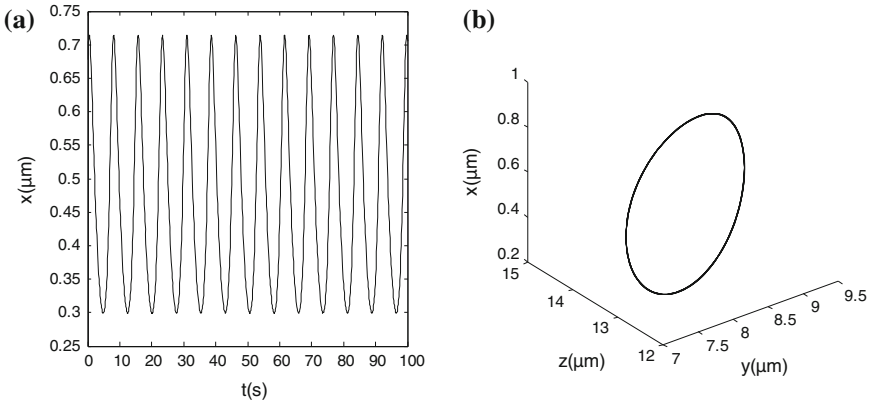
Fig. 2 a Temporal evolution for $r = 0.795$. **b** Phase portrait for $r = 0.795$

4 Numerical Simulation

Figure 1 is the bifurcation diagram of the equilibrium of system (2). Each point of the curve (solid line) means the stable equilibrium, and the dashed line indicates unstable equilibria. From Fig. 1 we can know that the equilibrium $O(r)$ undergoes bifurcation two times, marked as HB1 and HB2 of which the corresponding parameter values are $r_1 = 0.788038$ and $r_2 = 2.017361$ respectively. When $r < r_1$, there exists a stable equilibrium of system (2). With r increasing, the equilibrium loses its stability at HB1 and turns stable at HB2.

The curve as shown in Fig. 2a indicates temporal evolution of x of system (2) for r = 0.795. Figure 2b represents the limit cycle trajectory of system (2) in (z, y, x) phase space and its period is T $\approx$ 7.7 s. Fig. 3a shows the temporal evolution of x for r = 1.2 and Fig. 3b represents the limit cycle trajectory of system (2).

(a)

(b)
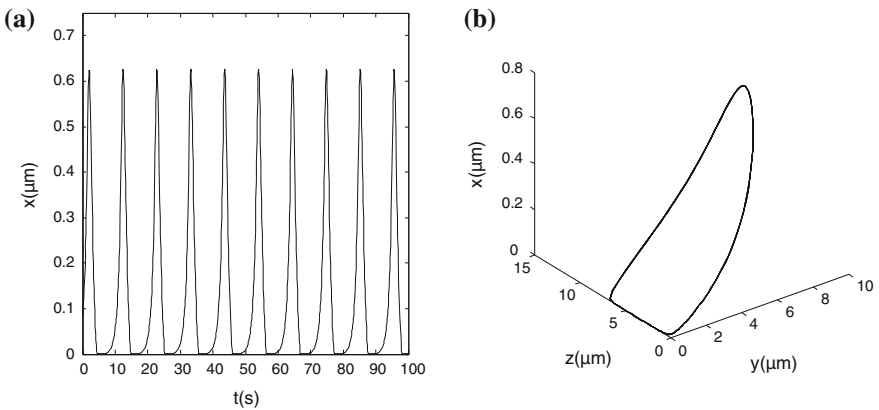
Fig. 3 a Temporal evolution for $r = 1.2$. **b** Phase portrait for $r = 1.5$
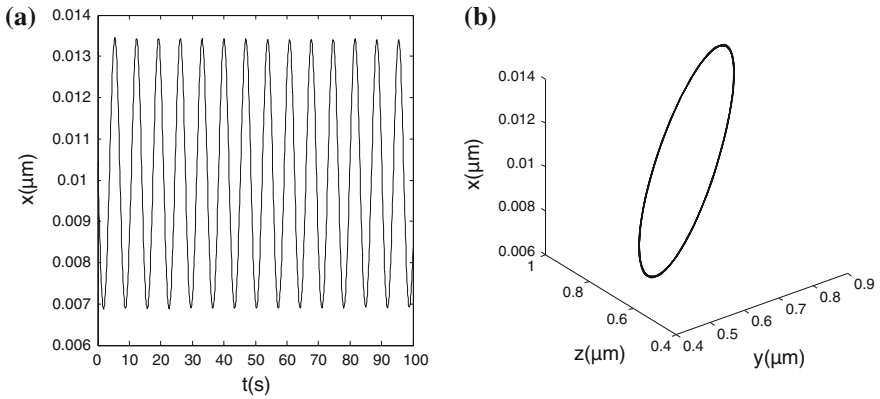
Fig. 4 **a** Temporal evolution for $r = 2$. **b** Phase portrait for $r = 2$

The period is T $\approx$ 10.5 s. When r = 1.2, there is a simple spiking with period 7 s as shown in Fig. 4a. The corresponding 3D phase portrait is Fig. 4b. From the results of numerical simulation we can know that the period of periodic solutions increase firstly, and the decrease with the increment of r when $0.788038 < r < 2.017361$.

Fig. 5 Temporal evolution for $r = 0.78$
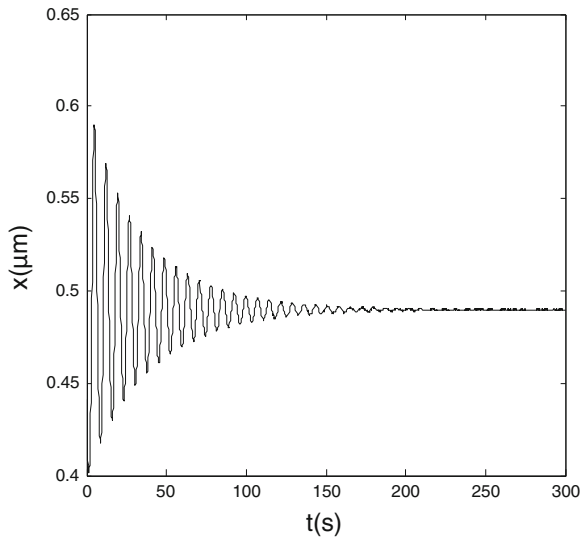
Fig. 6 Temporal evolution
for $r = 2.3$
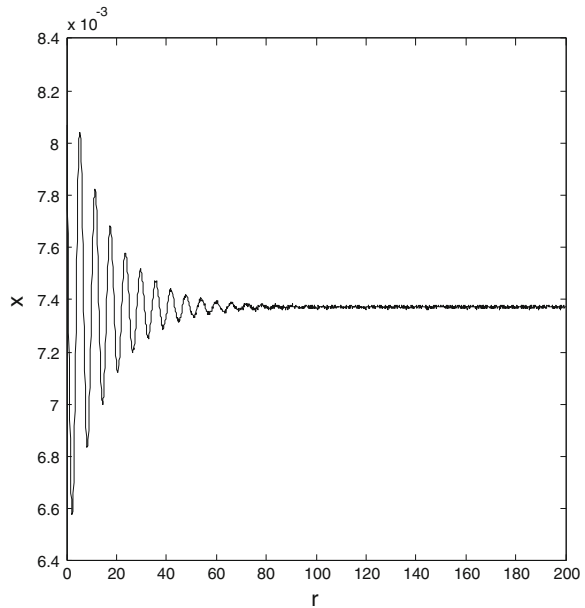
5 Conclusion

In this paper we analyze the existence, type and stability of equilibrium of Kummer-Olsen calcium oscillation model by using the qualitative theory of dynamical systems and stability theory.

For this sake we make a further study of the occurrence and disappearance of oscillations of this system based on the center manifold theory and bifurcation theory, which prove that the oscillatory phenomena are due to supercritical Hopf bifurcations of equilibria of this system.

Finally, with the help of Matlab software for numerical simulations, we verify that the results of theoretical analysis are effective Figs. 5 and 6.

References

1. Berridge MJ, Bootman MD, Roderick HL (2003) Calcium signaling: dynamics, homeostasis and remodeling. Nat Rev Mol Cell Biol 4:517–529
2. Woods NM, Kuthbertson KSR, Cobbold PH (1987) Agonist-induced oscillations in hepatocytes. Cells Calcium 8:79–100

3. Shen P, Larter R (2005) Chaos in intracellular calcium oscillations in a new model for non-excitable cells. Cell Calcium 17:225–232
4. Borghans JAM, Dupont G, Goldbeter A (1997) Complex intracellular calcium oscillations. A theoretical exploration of possible mechanisms. Biophys Chem 66:25–41
5. Grubelnk V, Larsen AZ, Kummer U, Olsen LF (2001) Mitochondria regulate the amplitude of simple and complex calcium oscillations. Biophys Chem 94:59–74
6. Meng P, Lu QS (2010) Dynamical effect of calcium pump on cytosolic calcium bursting oscillations with IP3 degradation. Chin Phys Lett 27:105–109
7. Wiggins S (1990) Introduction to applied nonlinear dynamical systems and chaos. Springer, Berlin
8. Errami H, Seiler WM, Eiswirth M, Weber A (2012) Computing Hopf bifurcations in chemical reaction networks using reaction coordinates. Lecture notes in computer science, vol 7442, pp 84–97
9. Budzko DA, Prokopenya AN (2012) Stability of equilibrium positions in the spatial circular restricted four-body problem. Lecture notes in computer science, vol 7442, pp 72–83

Study on Protein Structure Based on Reverse Hamilton Path Models

Xiaohong Shi

Abstract We present a new method for studying protein structure based on reverse Hamilton path models. A protein conformation with n sequences is changed to a weighted completely graph K_n. The reverse minimum Hamilton path in this graph matches the protein reverse sequence and n is positive proportion to length of reversed Hamilton path. The time of finding a reverse minimum Hamilton path grows quickly when n rise.

Keywords Protein conformation · A weighted graph · Reverse Hamilton path · Reverse Hamilton knot structure

1 Introduction

An understanding of structure leads to an understanding of function and mechanism of action. Therefore, structural knowledge of protein is a vital if we are to proceed to a complete understanding of life at the molecular level [1]. A protein is a long-chain molecule consisting of a backbone made up of amino acids connected sequentially via a peptide bond. For this reason, the chain is called a polypeptide chain. The number of units (amino acids) on the chain ranges from the order of 50–3,000. The amino acids are chosen from a set of 20 standard amino acids, and the sequence of the amino acids is called the primary structure. The first protein sequence was carried out by Sanger. By convention, the amino acid component retaining a free amine group is drawn at the left end (the N-terminus) of the peptide chain, and the amino acid retaining a free carboxylic acid is drawn on the right (the C-terminus) [2]. The concept of protein structural classes was proposed

X. Shi (✉)
College of Science, Xi'an Technological University, Xi'an 710032, China
e-mail: ishxh@163.com

by Levitt and Chothia on a visual inspection of polypeptide chain topologies in a set of 31 globular proteins [3]. They proposed ten structural classes, four principal and six small classes of protein structure. But the biological community follows the first four principal classes which are all α, all-β α/β and α + β. The *all* − α and *all* − β classes represent structures that consist of mainly α-helices and β-strands respectively [4]. The α/β and α + β classes contain both α-helices and β-strands which are mainly interspersed and segregated. Protein structural class information provides a key idea of their structure and also features related to the biological function.

Recently, Graph theoretical techniques have been presented for protein structure prediction and protein modeling [5]. We present a new method for studying protein structure based on reverse Hamilton path models of graph theory. A protein conformation is changed to a weighted completely graph K_n. The reverse minimum Hamilton path in this graph K_n matches the protein reverse sequence. The length of reversed Hamilton path is positive proportion to the number of the amino acids residue. The time of finding a reverse minimum Hamilton path grows quickly when the length of the sequence rise. The results show that reverse Hamilton knot structure is average appeared in different structural class, about 20 %. Results show that the reverse protein sequence is not a random coil but they have their mechanism and rule. Our results provide the theory data and new idea for protein folding and protein structure prediction.

2 Materials and Methods Heading

A weighted graph G is an ordered pair of disjoint sets (V, E, W) in which V and E are always finite. The V is the set of vertices and E is the set of edges. The weighted value W is related with two vertices and one edge. A protein structure is represented by a weighted graph in which the vertices of the graph represent the Cα atoms and the edges represent the distance between the Cα atoms. Then a protein with n residues length is changed to a weighted K_n. The Cα atoms of the residues are considered as vertices and the distance between the Cα atoms as edges. Each edge is weighted based on distance between the Cα coordinates of the two vertices. By convention, the protein sequence is started from N-terminus and end at the C-terminus. In this research we using reverse protein sequences and protein conformation to construct a weighted completely graph K_n. A reverse Hamilton path of the K_n is a path containing all the vertices once of the K_n. The reverse Minimal Hamilton Path (denoted *Hr-min)* is the shortest Hamilton path in a weighted graph. Assume that the first and the final vertex was known, this is a limited Hr-min Path conception. The number of reverse Hamilton path in this K_n graph is (n − 2)! Finding the Hr-min in a K_n graph is a NP-Complete problem.

Definition 1 *d* is adjacent distant such that *d* is the distant of two C_α atoms which are adjacent in the reverse protein sequence. φ_d is not adjacent distant such that φ_d

is the distant of two C_α atoms which are not adjacent in the reverse protein sequence. The average d is denoted by $\bar{d}$.

Definition 2 H^- is a reverse knot such that two C_α atoms have a reverse knot when find if aφ_d is small than d in the reverse Hamilton path. The protein structure is represented H^- structure classes. Otherwise the protein structure is *non- H^-* structure classes which denoted by φH^- structure classes.

Definition 3 H^+ is a knot such that two C_α atoms have a knot when find if aφ_d is small than d in the Hamilton path. The protein structure is represented H^+ structure classes. Otherwise the protein structure is *non-H^+* structure classes which denoted by φH^+ structure classes.

Dataset of protein chains is selected from the PDB (http://www.rcsb.org/pdb/). Each class is 18 and the total is 72 proteins. Classes contain $all - \alpha$, $all - \beta$, $\alpha + \beta$ and α/β structural classes. The Prim algorithm: Choose the first vertex (the C-terminus amino acid residue) and draw black, the others draw white. Choose an edge with minimal weight from the first vertex then the vertex is drawn black and other vertex is not. (If there are multiple edges with the same weight, any of them may be picked). Draw this edge black. Repeat this step unless all of the vertices are black. The *Hr-min* was made up of the black edges and black vertices. The average edge of the *Hr-min* in K_n is defined as *Hr-min* divided by $(n - 1)$ and is denoted by $< Hr\text{-}min >$.

The step in Matlab7.0

- Input: the adjacent matrix [a(i, j)]n * n of the connected weighted graph with n vertices.
- From the first vertex construct a candidate edges set E, in which every edge has two vertices with black and white.
- Construct the min-Hamilton path set H = [].
- Choose the minimal weight shortest edge in the candidate edges set E (vi, vj, wij), H = H + (vi,vj,wij), adjust the candidate edges set E.
- Repeat until H has $n - 2$ edges. Then: H = H + (v_j, v_n, w_{jn}),
- Compared natural H_{min} with the H, $H_{min} = \min (H_{min}, H)$
  Output: $< H_{min} > = H_{mn}/(n - 1)$ end.

3 Results and Discussion

We tested 72 protein structures based on the reverse Hamilton-path model of graph theory. The reverse minimum Hamilton path in this graph matches the protein reverse sequence. The time of finding a reverse minimum Hamilton path grows quickly when n rise.

The test result shown there are 14 H^- structure classes among the 72 protein structures with a rate of 19.4 %. H^- *knot* is found in the $all - \alpha$, $all - \beta$, $\alpha + \beta$

Table 1 The distribution rate of H^- knot and H^+ knot in the set for different structure classes

Classes	$all - \alpha$ (%)	$all - \beta$ (%)	α/β (%)	$\alpha + \beta$ (%)	Overall (%)
H^- knot	16.7	22.2	16.7	22.2	19.45
H^+ knot	22.2	11.1	22.2	22.2	19.43

and α/β structural classes. H^- knot can occurred in $all - \alpha$, $all - \beta$, $\alpha + \beta$ and α/β structural classes. The comparisons of the rate of the H^- knot and H^+ knot in different structure classes are listed in Table 1.

We tested 72 protein structures based on the reverse Hamilton-path model of graph theory. We find that amino acids component 12 H^- knot of the total 14 also be able to construct H^+ knot in the protein graph model that reached 85.7 % rate. Protein structure with a same H^- knot and H^+ knot in the test set distribution rate is 16.7 % in $all - \alpha$, 11.7 % in $all - \beta$, 16.7 % in α/β and 22.2 % in $\alpha + \beta$ structural classes. For example: 1ycc in the Hr-min model has a H^- knot which is component the 98 G and 10 L that two amino acids just correspond to the H^+ knot which is component the 6 G and 94 L. (as detail in finger 1) Fig. 1.

For another example 1 aac in the Hr-min model has a H^- knot which is component the 40 K and 47 A but they hasn't H^+ knot in the H-min model. Such H^- knot t is component of non-adjacent amino acids in sequence but display closer than the adjacent amino acids within the 3D structure [6]. Residue-residue contacts are known to play critical roles in maintaining the native fold of proteins and guiding protein folding. Our research of the H^- knot and H^+ knot is such residue–residue contacts and they mutual have a more potential energy so that may serve as an important model system for a deeper understanding of protein folding mechanisms and fold rate.

Fig. 1 Conformation of *1ycc* and H^+ knot

Finally, we present a new method for studying protein structure based on reverse Hamilton path models. A protein conformation with n sequences is changed to a weighted completely graph K_n. The reverse minimum Hamilton path in this graph matches the protein reverse sequence and n is positive proportion to length of reversed Hamilton path. The time of finding a reverse minimum Hamilton path grows quickly when n rise. The results may be of significant benefit in protein structure stability research and be useful to deep understand the protein folding mechanics.

4 Acknowledgement

References

1. Sanger F, Tuppy H (1951) The amino acid sequence of the phenylalanyl chain of insulin. Biochemical J 49(3):481–490
2. Anfinsen CB, Haber E, Sela M et al (1961) The kinetics of the formation of native ribo nuclease during oxidation of the reduced polypeptide chain. Proc Natl Acad Sci USA 47(9):1309–1314
3. Yan L,Sun Z (1999) protein molecular structure. Peking University Press, Beijing, pp 2–3
4. Chou KC (1995) A novel approach to predicting protein structure classes in a amino acid composition space. Proteins 21:319–344
5. Shi X, Liu W et al (2005) Study the protein structure prediction based on the graph theory. Biol Technol 15(5):89–92
6. Gromiha M, Thangakani M, Selvaraj S (2006) Fold-rate: prediction of protein folding rates from amino acid sequence. Nucleic Acids Res 34(2):70–74

Part II
Applications

Based on the Queuing Model of CAN Bus Simulation and Application

Jing Zhang and Tao Li

Abstract In this paper, a queuing model based on the CAN is established and its simulation and application are considered. The specific steps for, through the hypothesis model related parameters, make the whole model CAN achieve stable state. Through the calculation model of the probability of no data transmission, data in the model in the probability of transmission, data sending average queue goal number, average each data queue time, average each data in the model of the time and other related parameters modeling. By solving the queuing model, the data in the CAN bus to send on quantitative analysis, which CAN be used to CAN bus to transfer data distribution scheme optimization.

Keywords Queuing model · CAN bus · Data transmission · Probability calculation

1 Introduction

At present, the computer network, Field Control and other related content have made the Field bus Control System (FCS). Because CAN bus has simple structure, high reliability, the data communication node set being no-restriction, convenient

J. Zhang (✉)
College of Science, Jiujiang University, Jiujiang 332005, People's Republic of China
e-mail: 664723267@qq.com

T. Li
College of Computer and Information Engineering, Hohai University,
Nanjing 211100, People's Republic of China
e-mail: zhanglitaoj@sina.com

T. Li
College of Science, Anhui University of Science & Technology,
Huainan 232001, People's Republic of China

joining, communication having more strong real-time property, furthermore and communication media being no special requirements and easy to realize more advantages, such as the main structure, which led to the CAN bus has been widely used in industrial production of each related links [1, 2].

Queuing theory, called stochastic service system theory, it is mainly used to study statistical research through the service object and the arrival of the corresponding service time, from these data that data statistical rule of the relevant quantitative index is obtained, such as: the service object of waiting time, the service object queue length, desk busy period and idle period, etc. Then according to the related data statistics law, to improve the structure of the whole system or reconstruct, so we can make the whole system can meet the needs of the service object, and can make the system related index to achieve optimal. Queuing theory is the study of the system of the queue phenomena random regular discipline, and has been widely used in computer network resources sharing in the study of stochastic service system.

2 Queuing Model Overview

Generally speaking, a typical queuing model can be: one or more waiter queuing system, and mixed waiting in line model. From constitute a queuing phenomenon up analysis, although each factors each are not identical, but a queuing model is mainly composed of the following three parts place to form:

1. Reach mode: is used to describe the customer's source distribution, and according to the law of customer is how to service desk to line up of the law.
2. Queuing rules: refers to the service desk agree that began to queue, and desk agree according to what kind of queuing rules; And the corresponding is: customer is willing to on line, and according to what kind of rules in line. This rule summarized up, that is to say, the customer in the queue waiting for service, should according to what kind of rules in line. Usually are divided into the following three queue type:
3. Services: mainly is the study can provide services to the number of the front desk. The reception is to use the form of single service to provide services, or more than the front desk to provide service form. Especially in multiple service desk provides service, is to use series forms, or use the form of parallel. As the matching is, the customer to accept the service, service is to provide service customer batch or single service. And desk service time is to belong to what kind of probability statistical distribution, and the customer to accept the service time of the distribution is independent.

3 The Establishment of the System Model

Usually, according to the reality to structure reflect the whole system behavior motion mathematical model, including formal model and the formal model. In the classical queue model, usually with a letter to say a queuing system model, English letters interval is to use slash to interval. Regulation: the first English letters are used to represent the customer flow distribution types, the second English letters are used to represent desk are services provided service time distribution types, the third English letters used to represent the whole queuing model, desk allow customers to line up of the capacity, in addition, sometimes also used the fifth English letters customer source is provided in the number of customers.

This paper studies a set of customers (data) accept a parallel desk (namely CAN bus) services (transmission). Namely: a set of data to accept a parallel transmission of CAN bus. The many-to-many desk system model diagram, as shown in Fig. 1.

Next, in order to make model is simple, CAN the queue model simplified into a single queue model, that is, when a CAN bus exist idle time, then according to the corresponding transfer protocol, transmission under a set of data. The single for service system model diagram, as shown in Fig. 2.

3.1 Model Hypothesis

Because be in real life, there exist the following line structure and queuing rules, as shown in Figs. 3, 4.

Therefore, this paper assumes that the whole system model should be able to satisfy the following conditions:

1. In the whole system model, N of a completely the same CAN bus in the transmission;
2. In the whole system model, only a set of data in the transmission;
3. In the whole system model, data arrival is random arrival;
4. On the whole system model, the data in the CAN bus on transmission time is completely random;
5. In the whole system model, the data in the CAN bus on the order of the transmission is advanced to preach;
6. In the whole system model, the system can be arrived at stable state;

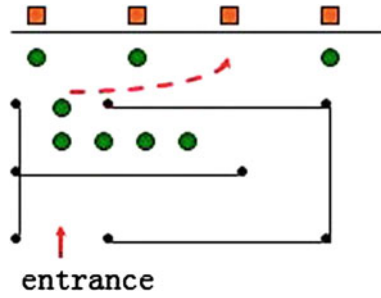
Fig. 1 Much more than team service desk
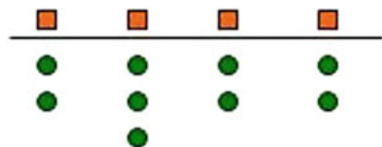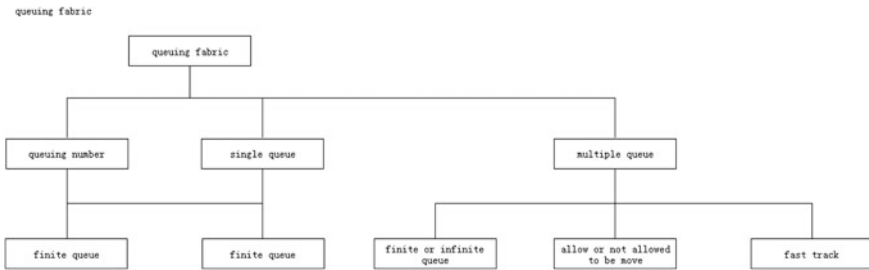
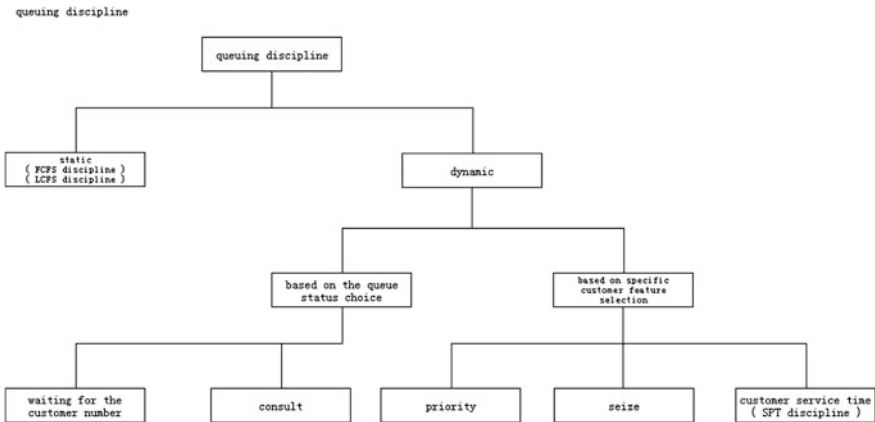**Fig. 2** Single team many desk

Fig. 3 Queue structure

Fig. 4 Queuing rules

7. In the whole system model, not limited to the number of data in the system number;

8. In the whole system model, the data transmission to start the number of unlimited number;

9. In the whole system model, as long as the data has been on the CAN bus, it is immediately transmission, no any limits.

   This paper used the related symbol for:

λ Unit of time to enter the CAN bus for transmission average data number's number

μ Each CAN bus data transmission speed

   So in order to make the whole system model to achieve stable state, it must ensure that: $N \cdot \mu \succ \lambda$.

3.2 Establishment of the Model

When all the a CAN have an idle period condition, that is the whole system model, countless according to the probability of transmission should be:

$$P_0 = 1 \left/ \sum_{n=0}^{N-1} \left(\frac{\lambda}{\mu}\right) \right/ n! + \left(\frac{\lambda}{\mu}\right)^N \cdot \mu \left/ (N-1)! \cdot (N.\mu - \lambda) \right.$$

If a data in the whole system model of probability should be:

$$P_n = \begin{cases} [(\lambda/\mu)^n / n!] \cdot p_0 & 0 \leq n \leq N \\ [(\lambda/\mu)^n / N! \cdot N^{n-N}] \cdot p_0 & n \succ N \end{cases}$$

In the whole system model, the relevant system characteristic parameters for:

1. The average queue waiting for CAN bus transmission data number's number is:

$$L_q = \left[(\lambda/\mu)^N \cdot \lambda \cdot \mu/(N-1)! \cdot (N \cdot \mu - \lambda)^2 \right] \cdot p_0$$

2. The average in the system the data in the model number's number is:

$$L = L_q + (\lambda/\mu)$$

3. The average data in the system model of the queue time for:

$$W_q = L_q/\lambda$$

4. The average data in the system model of the residence time for:

$$W = W_q + (1/\mu)$$

4 Model Application Examples

Assume that the existing a system, this system in the number of CAN bus for $N = 6$, in this system, each group of data is random coming, average every millisecond for 80. Each data in the transmission of CAN bus and eventually to the average amount of time 3 ms.

In the whole system model, each group of the average data to arrive speed for: $\lambda = 80\uparrow /ms$, each CAN bus average transmission speed for: $\mu = 20\uparrow /ms$, $N = 6$. So the problems can be attributed to $M/M/3$ queuing problem, then calculating the system model, the probability of a data for: n.

P0	P1	P2	P3	P4	……
0.017	0.067	0.133	0.177	0.177	……

Wait in line average data number's number should be:

$$L_q = \left[(\lambda/\mu)^N \ \cdot \ \lambda \ \cdot \ \mu/(N-1)! \ \cdot \ (N \cdot \mu - \lambda)^2 \right] \ \cdot \ p_0 = 0.580$$

In the whole system model, the average number of data required for:

$$L = L_q + (\lambda/\mu) = 4.580$$

In the whole system model, the average data transmission time waiting for:

$$W_q = L_q/\lambda = 0.435 \ (\text{min})$$

In the whole system model, the average data in the whole system model of sojourn time for:

$$W = W_q + (1/\mu) = 3.435 \ (\text{min})$$

5 Model Simulation

Now, general performs queuing theory research, mainly from the following two aspects to: (1) The Markov stochastic process; (2) Queuing model and its related performance analysis. In the two aspects of the study, queuing theory shows phenomenon is conceptual strong, formula theorem, and difficult to understand. At present, the problem has a very good solution. A can be queuing theory and related system performance analysis is an effective tool Java Modeling Tools (hereinafter referred to as the JMT) [3], can for scientific research service, the related research progress can view references [4–9].
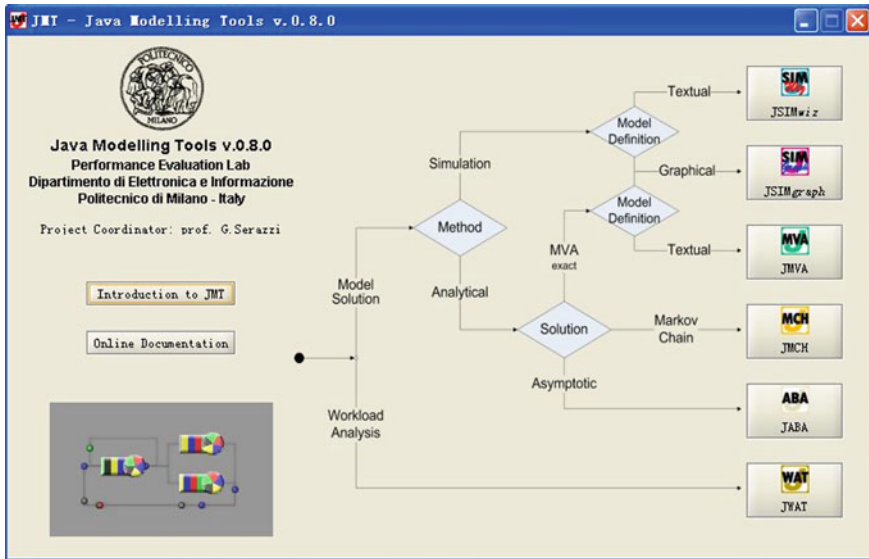
Fig. 5 JMT simulation tool selection interface

Next, this paper, by using for JMT to model simulation (Fig. 5).

In the face of the examples under simulation, in the whole system model, each group of the average data to arrive speed for: $\lambda = 80/ms$, each CAN bus average transmission speed for: $\mu = 20/ms$, $N = 4$. Through the derivation average queue length, utilization, the whole system response time, and the whole system input output of basic mathematical formula, and then establish JSIMgraph as below model, through the establishment of the model, parameter setting and the solving, to complete the model simulation (Fig 6).

Based on the model simulation, when the whole system is in a state of balance, right now, the system of the inflow and outflow is equal, so that we can work out the model of analytic solution (Fig 7).

In the example of the simulation process, and may, according to the average data in the whole system model of the residence time, namely data in CAN bus began to transmission to complete transmission time. CAN assume that each CAN bus for data transmission speed, the length of the CAN bus, the data in CAN bus transmission time CAN be calculated, should be: $t = m/v$. So, in order to make the whole system of data transmission time shortened, it must meet the requirements: t <= w. Therefore, in order to make the whole system of data transmission time shortened, can shorten. This with our common sense is consistent; namely, shorten the data in the whole system model of the average residence time.
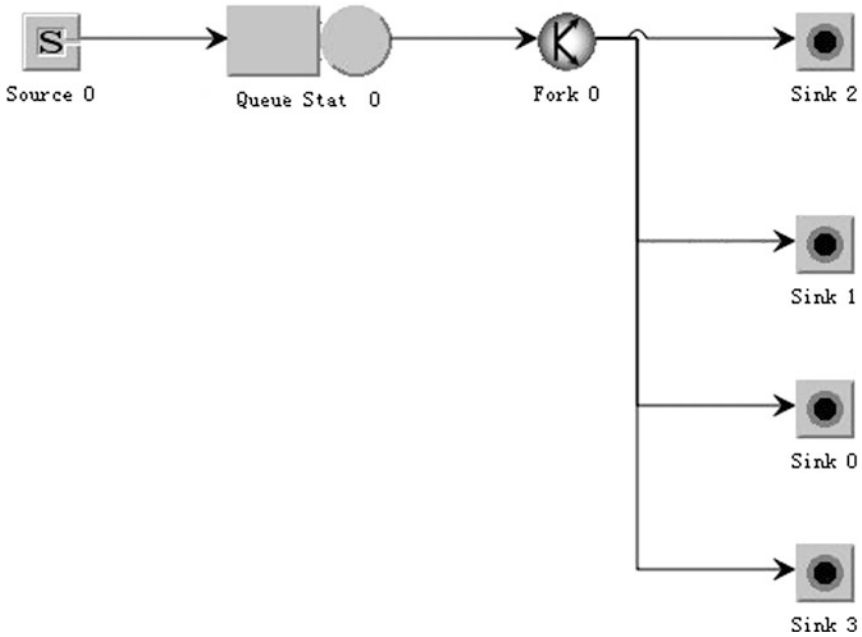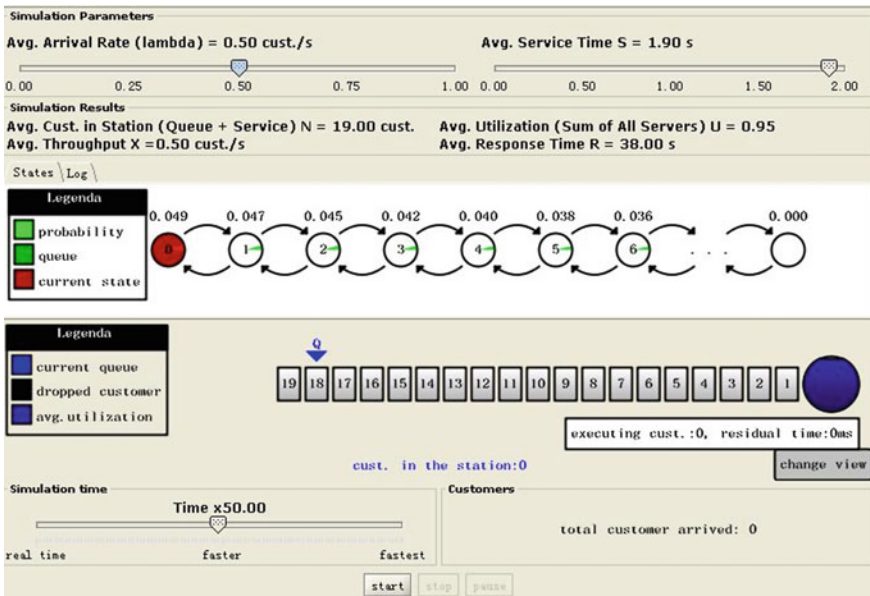
Fig. 6 JMT simulation interface

Fig. 7 Dynamic demonstration of diagram

6 Conclusion

From the above model CAN, it is that by using the queuing model of CAN bus CAN be very good simulation data transmission process, and CAN be the data transmission process do quantitative description, CAN also be for CAN bus configuration provide guidance, thus CAN improve the whole system model efficiency of the system.

References

1. Ding N, Ma F (2003) Monitoring system and filedbus. China University of Mining and Technology Press, Xuzhou
2. Lei H (2002) The distributed SCADA monitoring and control system based CAN bus in coal mine. Taiyuan University of Technology Press, Taiyuan, Shanxi
3. Serazzi G (Ed) (2008) Performance evaluation modelling with JMT: learning by examples Politecnico di Milano. DEI, TR, 2008.09, pp 366
4. Casale G, Serazzi G (2011) Quantitative system evaluation with java modelling tools. In: Proceedings of 2nd ACM/SPEC international conference on performance engineering, Karlsruhe, (ICPE), Germany, March (2011)
5. Bertoli M, Casale G, Serazzi G (2009) JMT: performance engineering tools for system modeling. ACM SIGMETRICS Perform Eval Rev 36(4):10–15
6. Bertoli M, Casale G, Serazzi G (2009) User-friendly approach to capacity planning studies with java modelling tools. In: Proceedings of international ICST conference on simulation tools and techniques, SIMUTools 2009, ACM press, Rome, Italy
7. Bertoli M, Casale G, Serazzi G (2007) The JMT simulator for performance evaluation of non-product-form queueing networks. In: Proceedings of SCS annual simulation symposium 2007, IEEE Press, Norfolk, VA, US, pp 3–10
8. Bertoli M, Casale G, Serazzi G (2007) An overview of the JMT queueing network simulator. Politecnico di Milano—DEI, TR 2007.2
9. Bertoli M, Casale G, Serazzi G (2006) Java modelling tools: an open source suite for queueing network modelling and workload analysis. In: Proceedings of QEST 2006 conference, IEEE Press, Riverside, US, pp 119–120, Sep 2006

Author Biographies

Jing Zhang born in 1979, MS degree, Jiujiang university lecture. Her main research interests include information security and cryptography theory and technology.

Tao Li born in 1979, Ph.D. Candidate, Hohai University student. His main research interests include information security and cryptography theory and technology.

Hierarchical Modeling Fault-Error-Failure Dependencies for Cyber-Physical Systems

Shixi Liu, Xiaojing Hu and Jingming Wang

Abstract While cyber-physical system (CPS) has grown in size and complexity, the existing single node failure of system are confronting some serious challenges. The strong coupling of software and physical processes in the emerging field motivates the development of new methods to respond to failure in both the cyber and physical domains. This paper presents formalized definitions for CPS from the view of hierarchical level model. The relationships between faults, errors and failures in the resource, service and process level are introduced. The application of stochastic Petri nets (SPN) on fault-error-failure dependencies Model in CPS is presented. Based on the model analyzing, the Fault-Error-Failure dependency relationships between resource, service and process are presented. The simulation results show the chances of resource faults are larger than the service error; the effects of a non-detected error perceived from resource level are less than that for the service level.

Keywords Fault · Error · Failure · Cyber-physical system · Stochastic petri nets

1 Introduction

The strong coupling of software and physical processes in the emerging field of cyber-physical systems motivates the development of new methods to respond to failures in both the cyber and physical domains. With the inherent complexity of CPS it is extremely difficult to precisely and deterministically describe the physical, technological, and environmental factors and interactions that provoke CPS malfunctioning. Achieving fault tolerance requires safety analysis that is best performed beginning with the earliest design stages. Part of such design analysis is the so-called fault propagation analysis of system model [1].

S. Liu (✉) · X. Hu · J. Wang
Computer Science and Technology Department, Chuzhou University, Chuzhou, China
e-mail: liusxchuz@163.com

Since the cyber-physical systems are composed of software interacting with the physical world, many classes of faults exist. On the cyber side, there are timing failures of real-time programs and operating systems, in addition to crash failures, and simply software bugs. On the physical side, there are actuator, control surface, and sensor failures, aside from robustness given the potential operating environments of a system. Between these two worlds is the potential for communication failures, such as message drops and omissions, or worse, adversarial man-in-the-middle attacks perhaps culminating in Byzantine failures [2].

When the delivered service of a CPS deviates from fulfilling the system intended function, the CPS has a failure. A failure is due to a deviation from the correct state of the CPS, known as error. Such a deviation is due to a given cause, for instance related to the physical state of the system, or to a bad system design. This cause is called a fault. It generically refers to fault, errors, and failures as the "FEF elements" [2].

The dependency relationships between faults, errors, and failures followed in this paper is illustrated in Fig. 1. Internal propagation is caused by the computation process of the faulty cyber system while the external propagation from cyber system to physical system, which receives control operation from cyber system, occurs when, through internal error propagation, an error reaches the cyber-physical interface. Then the control operation delivered by cyber system to physical system becomes incorrect provoking the failure of physical system that appears as an external fault to physical system and propagates the error into physical system.

The remainder of this paper is arranged as follows: Sect. 2 introduces the related work; Sect. 3 introduces typical architecture of Cyber-Physical Home Control System architecture, and presents 2 kinds of definitions of Cyber-physical system from the view of resource, service and process; Sect. 4 talks about the taxonomy of fault, error, failure, and takes advantage of SPN to model fault-error-failure dependencies in CPS, finally the simulation results are showed in this section; The last section summarizes the conclusion and claims the future work.
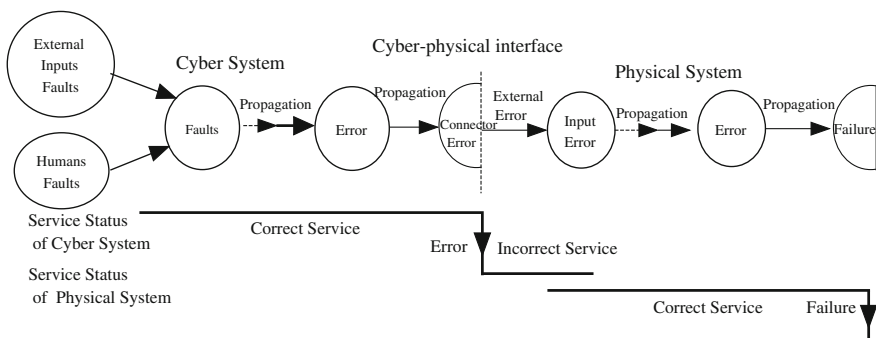

Fig. 1 Relationships between faults, errors and failures

2 Related Work

An inductive method to study fault propagation is the Failure Propagation and Transformation Analysis (FPTA) [3]. A failure logic for each component allows for automated analysis of the system. Here, the failure logic depends only on the inputs, which is appropriate for modeling software failures, but cannot explicitly use failure states of the component itself.

Taylor Johnson points out that Potential Directions in Cyber-Physical Systems are the handling of failures [1]. The complexity of this problem, solely in the classes of faults, is great. In addition to exploring the applicability of the existing literature and the new proposed methods to CPS, people must also utilize first principles such as abstraction to deal with this complexity, as otherwise that will wind up over designing these new systems. Some interesting questions are, given a system and a model, when is it possible to detect failures; Can time be bounded to detection; how do we minimize time to detection.

Linda Briesemeister presents an implementation of a probabilistic logic model, which allows for reasoning via symbolic evaluation as well as numeric evaluation to perform a quantitative fault analysis [4]. Formal methods and verification will provide useful tools for solving these problems. Sha motivated in [5] the need for formal methods in detecting and mitigating faults in what are now termed CPS, and more recently again as there are countless further directions to explore [6]. While thus far they have primarily considered mitigating faults to avoid violation of safety properties, they propose also a study of optimizing liveness properties in addition to maintaining safety.

3 CPS Definitions and Architecture

Architectures typically represent systems at a higher level than simulation models, which represent the details of a particular system implementation. Although architectural modeling has been used in specific domains to incorporate physical elements as components, there is currently no way of treating cyber and physical elements equally in a general fashion. This is because the components and connectors in software architectural styles are inadequate for representing the types of physical components found in CPS and their interactions with cyber entities.

Figure 2 shows the entire architecture of the Cyber-Physical Home Control System [7]. The Cyber-Physical Home Control System proposed contains three layers: the Physical Layer, the protocols and profiles for users control of home appliances; the Service Layer, the conversion and management of Physical Layer signals and Application Layer services, through device drivers in the OSGi Bundle, control signals are in packet formats, which conform to the specific protocols of the home appliances of the Physical Layer; the Application Layer, all the services provided by physical equipment can be converted to the OSGi service
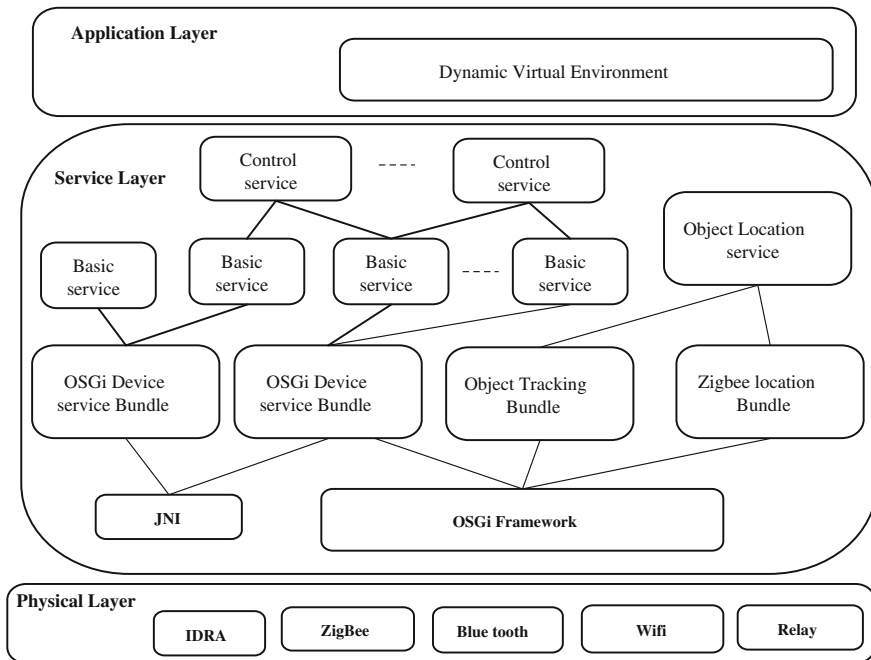
Fig. 2 Cyber-Physical Home Control System architecture

model for users, through the Service Layer. Details of the various components of the system architecture proposed are described as in reference [6].

This section presents extensions that allow fault, error, and failure elements to be modeled together in a CPS architectural style. A formalized definition of Cyber-physical system free of faults, errors, failures is presented by Definition 1.

Definition 1 $CPS = \{R, S, P\}$, $R = \{r_1, r_2, \ldots, r_m\}$; $S = \{s_1, s_2, \ldots, s_n\}$; $P = \{p_1, p_2, \ldots, p_t\}$.

The CPS is organized into three levels: resources, services and processes represented by R, S, P respectively. Resources are at the bottom level, and they provide operations for the services, where a service is basically a complex pattern of use of the resources. Services are then requested by the application model placed at the highest level, called process level.

A formalized definition of Cyber-physical system with faults, errors, failures is presented by Definition 2.

Definition 2 $CPS = \{R, S, P\}$, $R = \{R', FT\}$, $S = \{S', E\}$, $P = \{P', Fn\}$; $R' = \{r_1, r_2, \ldots, r_m\}$, $FT = \{ft_1, ft_2, \ldots, ft_m,\}$; $S' = \{s_1, s_2, \ldots, s_n\}$, $E = \{e_1, e_2, \ldots, e_n\}$; $P' = \{p_1, p_2, \ldots, p_t\}$, $Fn = \{fn_1, fn_2, \ldots, fn_t, \}$.
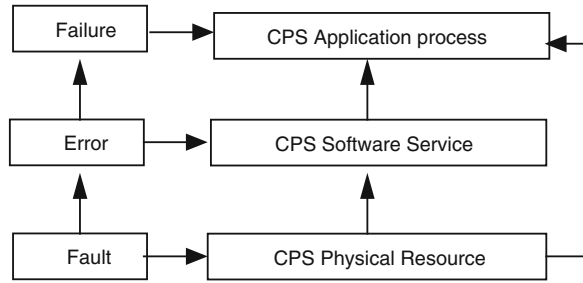
It assumes that there are a number of models for faults, errors, and failures, called FT-$_j$, E-$_h$, and F-$_n$, respectively, and it provides a schematic view of how they are organized in three levels: fault models are placed at the resource level, error models at the service level and failure models at process level. We assume that faults affect only the behavior of the resources of the CPS, errors are perceived at the service level, while failure are a concern of CPS Components, and therefore of the process level.

Figure 3 [2] shows the relationships between faults, errors and failures as well as their propagation among the cyber-physical system entities. Once customized on a specific application, it shows which faults provoke which errors, and which errors provoke a failure, that is to say a deviation from the function delivered by the cyber-physical system. Figure 3 also connects each type of fault, error and failure with the corresponding cyber-physical system entity affected by it, so a fault may affect a resource, an error may affect a service performed on one or more faulty resources, and a failure may affect the whole cyber-physical system if errors are not recovered in due time.

4 The SPN Model of Failure-Error-Fault Dependencies in CPS

4.1 Introduction to Stochastic Petri Nets

Stochastic Petri nets are a graphical and mathematical modeling tool applicable to many systems. They are a promising tool for describing and studying information processing systems that are characterized as being concurrent, asynchronous, distributed, parallel, nondeterministic [7]. SPN model expresses a complex system in the form of graph, and its structure elements include Place, Transition and Arc.

Definition 3 [7] Stochastic Petri Net is a quadruple (P, T; F, λ) where,

(1) P is a finite set of places;
(2) T is a finite set of transitions ($P \cap T \neq \varphi$);
(3) $F \subseteq (P \times T) \cup (T \times P)$ is a set of arcs;

(4) $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ is a set of firing rates of transitions set.

The state space of SPN will increase exponentially along with problems becoming more complex. However, it makes the isomorphic Markov Chain of SPN hard to solve. The presentation of GSPN (Generalized Stochastic Petri net) offers an effective approach to solve the problem of state space explosion. To the detailed, as in Murata [7].

4.2 The SPN Model of Fault-Error-Failure Dependencies in CPS

Figure 4 is the SPN model of Cyber-Physical Home Control System which is organized into three levels: resources, services and processes learning from reference [8]. Resources are at the bottom level, and they provide operations for the services, where a service is basically a complex pattern of use of the resources. Services are then requested by the application model placed at the highest level, called process level. The top part of Fig. 4 depicts a skeleton of the process model that uses services. Suppose process of CPS carried out by resource components and service components. Then, P_{up} and $P_{failure}$ places' markings which mean the up state and down state of process depend upon the markings of the resource and service components models.

The activation rate of a resource fault is λ_h (T_{r1}), and of a service error is λ_s (T_{r6}); R_{fa} and S_{ea} means fault activated and error activated respectively. The probability that a resource fault is temporary is t (t_{r1}). Such faults will disappear with rate ε (T_{r2}); A permanent resource fault (resp. service) is detected by the fault-tolerance mechanisms with probability d_h (resp. d_s for service faults). The detection rate is δ_h (T_{r3}) for the resource, and δ_s (T_{r7}) for the service; place R_{fault} and S_{error} means the resource and service of CPS is in fault state and error state respectively. The effects of a non-detected error are perceived with rate π_R (T_{r4}) for the resource, and rate π_S (T_{r8}) for the service; Place R_{fnd} and S_{end} means fault does not detected and so doesn't error respectively.

Faults detected in the resource component require its repair: repair rate is $\mu(T_{r5})$; permanent errors in the service may necessitate only a reset, place S_{rs} means service reset. The reset rate is $\rho(T_{r9})$ and the probability that an error induced by the activation of a permanent service error disappears with a reset is $\gamma(tr_8)$. If the error does not disappear with the service reset, a re-installation of the service is done, place S_{ri} means service reinstallation. The service's re-installation rate is $\sigma(T_{r10})$. In the following, we will talk about dependency analysis between faults, errors, and failures.
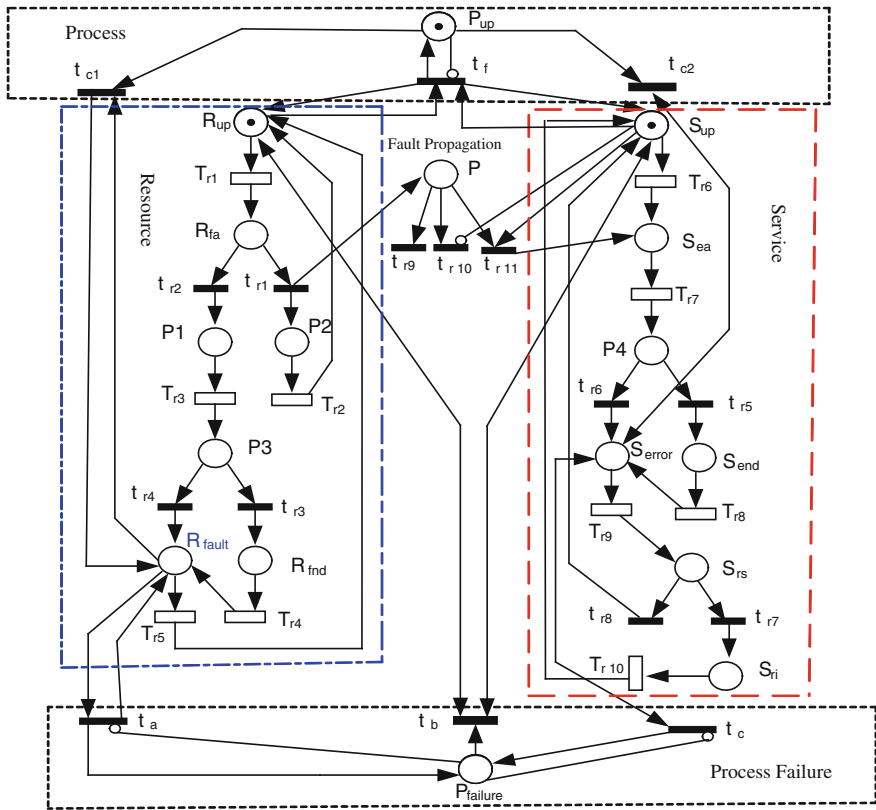
Fig. 4 The SPN model of Cyber-Physical Home Control System

4.3 Simulation Results

The average throughput E[Ta] for transition a is defined as $E[Ta] = \sum_{i \in R(a)} p(i) *$
$\rho(t, i)$ where R(t) is the subset of reachable markings that enable Transition t, p(i) is the probability of marking i, and $\rho(a, i)$ is the rate of transition t in marking i. For each timed transition, Pr[enabled] is the probability which it is enabled. Pr[nonempty] means the probability that a place is not empty, and Av[tokens] means average number of tokens in a place.

Here, we assume that $\lambda_h = 0.25$, $\lambda_s = 0.01$, $\varepsilon = 0.23$, $d_h = 0.7$, $d_s = 0.66$, $\delta_h = 0.25$, $\delta_s = 0.1$, $\pi_R = 0.24$, $\pi_S = 0.0.26$, $\mu = 0.05$ $\rho = 0.02$, $p = 0.01$, $\gamma = 0.8$, $\sigma = 0.01$. We take advantage of Stochastic Petri Net Package (SPNP) developed by Duke University to run the SPN model in Fig. 4. The average throughput of Pr[enabled], Pr[nonempty], and Av[tokens] are solved by the SPNP. According to the simulation results, the enabled probabilities of transition Tr1 and

Tr6 are 3.423423424351e-001 and 2.352352351444e-001 respectively; the average throughputs of them are 8.558558560878e-002 and 5.880880878609e-002. That means the chances of resource faults are larger than that of service error. It is maybe the CPS elements are in harsh environment which makes the faults activated frequently. The enabled probabilities of transition Tr4 and Tr8 are 7.558558555380e-002 and 9.409409411586e-002; the average throughputs of them are 1.139639638845e-002 and 2.352352352897e-002. That means the chance of effects of a non-detected error is perceived for the resource level are less than that for the service level. The cause is that the repair of software for service level is much easier than that of resource level. The enabled probabilities of transition Tr5 and Tr10 are 1.711711711152e-001 and 1.176176176465e-001; the average throughputs of them are 4.279279277880e-002 and 2.940440441163e-002. That means the repair of errors detected in the resource level is larger than that of re-installation of the service level when error does not disappear with the service reset.

5 Conclusion

In CPS, people are interested in characterization faults and how faults manifest as failures that affect overall dependability goals of the system. Formal methods and verification will provide useful tools for solving these problems. The modeling approach presented in this paper presents a generally-applicable process for fault-error-failure dependencies analysis in CPS based on GSPN. The simulation results show enabled probability and average throughput of timed transitions, and the average tokens and nonempty probability of places are presented too. The results show that CPS elements are in harsh environment which makes the faults activated frequently. The up state of process of CPS depends on not only the resource level but also the service level. Next work is to analyze the dependability of cyber-physical system using SPN.

References

1. Johnson TT, Mitra S (2009) Handling failures in cyber-physical systems: potential directions. Presented at RTSS 2009 PhD student forum
2. Bernardi S (2003) Building Stochastic Petri Net models for the verification of complex software systems. PhD thesis, Dipartimento di Informatica, Universit'a di Torino

3. Ge X, Paige R, McDermid J (2010) Analysing system failure behaviours with PRISM. In: International conference on secure software integration and reliability improvement companion, pp 130–136
4. Briesemeister L, Denker G, Elenius D, Mason I, Varadarajan S, Bhatt D, Hall B, Madl G, Steiner W (2011) Quantitative fault propagation analysis for networked cyber-physical systems. In: 2nd analytic virtual integration of cyber-physical systems workshop (AVICPS)
5. Sha L (2001) Using simplicity to control complexity. IEEE Softw 18(4):20–28
6. Laia C-F, Maa Y-W, Changa S-Y, Chaob H-C, Huanga Y-M (2011) OSGi-based services architecture for Cyber-Physical Home Control Systems. Comput Commun 34(2):184–191
7. Murata T (1989) Petri nets: properties, analysis, and applications [J]. Proc of the IEEE 77(4):541–580
8. Betous-Almeida C, Kanoun K (2001) Dependability evaluation: from functional to structural modelling. In: Proceedings of the 20th international conference on computer safety, reliability and security (SAFECOMP 2001), LNCS 2187, pp 227–237

A Comparative Study of State Transition Algorithm with Harmony Search and Artificial Bee Colony

Xiaojun Zhou, David Yang Gao and Chunhua Yang

Abstract We focus on a comparative study of three recently developed nature-inspired optimization algorithms, including state transition algorithm, harmony search and artificial bee colony. Their core mechanisms are introduced and their similarities and differences are described. Then, a suit of 27 well-known benchmark problems are used to investigate the performance of these algorithms and finally we discuss their general applicability with respect to the structure of optimization problems.

Keywords State transition algorithm · Harmony search · Artificial bee colony

1 Introduction

Existing natural phenomena, such as natural selection and survival of the fittest (genetic algorithm), natural annealing process in metallurgy (simulated annealing), foraging behavior of real ant colonies (ant colony optimization), and social behavior of bird flocks and fish schools (particle swarm optimization) have inspired researchers to develop algorithms for optimization problems. These nature-inspired algorithms have received considerable attention due to their strong

X. Zhou (✉) · D. Y. Gao
School of Science, Information Technology and Engineering, University of Ballarat,
Ballarat, VIC 3350, Australia
e-mail: tiezhongyu2010@gmail.com

D. Y. Gao
e-mail: d.gao@ballarat.edu.au

X. Zhou · C. Yang
School of Information Science and Engineering, Central South University,
Changsha 410083, China
e-mail: ychh@csu.edu.cn

adaptability and easy implementation. Inspired by the improvisation process of musicians and foraging behavior of real honeybees, harmony search (HS) [1–3] and artificial bee colony (ABC) [4, 5] have been proposed respectively in recent few years. At the same time, in terms of the concepts of state and state transition, a new heuristic random search algorithm named state transition algorithm (STA) has been introduced in order to probe into classical and intelligent optimization algorithms [6–9]. In this study, we focus on a comparative study of state transition algorithm with harmony search and artificial bee colony in their standard versions.

2 Three Stochastic Algorithms

In this section, we give a brief description of the three stochastic algorithms with respect to their mechanisms, and the similarities and differences are also discussed.

2.1 Harmony Search

In HS, there exist three possible choices to generate a new piece of music: (1) select a note stored in harmony memory at a probability of HMCR (harmony memory considerate rate); (2) adjust the pitch slightly at a probability of PAR (pitch adjusting rate); (3) compose any pitch randomly within bounds. The pitch is adjusted by

$$x_{new} = x_{old} + (2rand - 1) * b$$

where, $rand$ is a random number from [0,1], and b is the pitch bandwidth.

2.2 Artificial Bee Colony

In ABC, the colony of artificial bees contains three groups of bees: (1) employed bees, going to the food source visited previously; (2) onlookers, making decision to choose a food source; (3) scouts, carrying out random search. A new position is produced by

$$v_{ij} = x_{ij} + \phi_{ij}(x_{ij} - x_{kj})$$

where, i is the index of ith food position, j is the jth component of a position, φ_{ij} is a random number in $[-1,1]$, k is a different index from i, and j, k are created randomly.

An artificial onlooker bee chooses a food source depending on a probability by

$$p_i = \frac{\text{fit}_i}{\sum\limits_{n=1}^{SN} \text{fit}_n}$$

where, fit_i is the fitness value of the position i, SN is the number of food sources.

2.3 State Transition Algorithm

The unified framework of STA is described as follows

$$\begin{cases} x_{k+1} = A_k x_k + B_k u_k \\ y_{k+1} = f(x_{k+1}) \end{cases}$$

where, A_k, B_k are state transition matrix, u_k is the function of state x_k and historical states, and there are four special geometrical operators defined by

1. Rotation transformation

$$x_{k+1} = x_k + \alpha \frac{1}{n\|x_k\|_2} R_r x_k,$$

where, α is a positive constant, R_r is a random matrix with its entries from $[-1,1]$.

2. Translation transformation

$$x_{k+1} = x_k + \beta R_t \frac{x_k - x_{k-1}}{\|x_k - x_{k-1}\|_2},$$

where, β is a positive constant, R_t is a random variable from $[0,1]$.

3. Expansion transformation

$$x_{k+1} = x_k + \gamma R_e x_k,$$

where, γ is a positive constant, R_e is a random diagonal matrix with its entries obeying the standard norm distribution.

4. Axesion transformation

$$x_{k+1} = x_k + \delta R_a x_k$$

where, δ is a positive constant, R_a is a random diagonal matrix with its entries obeying the standard norm distribution and only one random position having nonzero value.

2.4 Similarities and Difference

There are two main similarities among the three algorithms in the discussed versions: Firstly, a new solution is created randomly, and they are all stochastic algorithms. Second, "greedy criterion" is adopted to evaluate a solution, and it is different from simulated annealing, in which, a bad solution is accepted in probability.

The differences between STA and other two algorithms are: (1) both HS and ABC focus on updating each component of a solution, while STA treats a solution in whole for update except the axesion transformation; (2) the comparing STA is individual-based, while both HS and ABC are population-based; (3) the mutant operators are different in three algorithms; (4) in HS, there is a probability in choosing an update, while in STA, the updating procedures are determined; (5) in ABC, choosing a food source depending on a probability associated with the fitness, while in STA, a candidate solution with better fitness is preferred; (6) in ABC, the fitness is standardized, while in STA, the fitness is based on objective function.

3 Experimental Results

All these benchmark instances are taken from [10]. In our experiments, we use the codes of standard HS and ABC from [11, 12], and the STA is from Zhou et al. [7]. The size of the population is 10, and the maximum iterations are 1e3, 2e3, 4e3, 1e4, 5e4, and 1e5 for $n = 2, 3, 4, 10, (20, 24, 25)$ and 30, respectively. For each benchmark instance, the initial population is the same for three algorithms at each run, and 20 runs are performed for each algorithm. Statistics like mean, std (standard deviation), and Wilcoxon rank sum test are used to evaluate algorithms.

3.1 Benchmark Instances

The details of the benchmark instances are given as follows.
   Ackley function

$$f_1(x) = -20 \exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}\right) - \exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)\right) + 20 + e, \; -15 \le x_i \le 30$$

Beale function

$$f_2(x) = (1.5 - x_1 + x_1 x_2)^2 + (2.25 - x_1 + x_1 x_2^2)^2$$
$$+ (2.625 - x_1 + x_1 x_2^3)^2, \ -4.5 \le x_i \le 4.5$$

Bohachevsky Function

$$f_3(x) = x_1^2 + 2x_2^2 - 0.3\cos(3\pi x_1) - 0.4\cos(4\pi x_2) + 0.7, \ -100 \le x_i \le 100$$

Booth Function $f_4(x) = (x_1 + 2x_2 - 7)^2 + (2x_1 + x_2 - 5)^2, \ -10 \le x_i \le 10$

Branin Function

$$f_5(x) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos(x_1) + 10, \ -5 \le x_1 \le 10, \ 0 \le x_2 \le 15$$

Colville Function

$$f_6(x) = 100(x_1^2 - x_2)^2 + (x_1 - 1)^2 + (x_3 - 1)^2 + 90(x_3^2 - x_4)^2$$
$$+ 10.1\left((x_2 - 1)^2 + (x_4 - 1)^2\right)$$
$$+ 19.8(x_2 - 1)(x_4 - 1), \ -10 \le x_i \le 10$$

Dixon and Price Function $f_7(x) = (x_1 - 1)^2 + \sum_{i=2}^{n} i(2x_i^2 - x_{i-1})^2, \ -10 \le x_i \le 10$

Easom Function $f_8(x) = -\cos(x_1)\cos(x_2)\exp(-(x_1 - \pi)^2 - (x_2 - \pi)^2),$ $-100 \le x_i \le 100$

Goldstein and Price Function

$$f_9(x) = \left(1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 13x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)\right)$$
$$\times \left(30 + (2x_1 - 3x_2)^2(18 - 32x_1 + 12x_1^2 - 48x_2 - 36x_1x_2 + 27x_2^2)\right),$$
$$-2 \le x_i \le 2$$

Griewank Function $f_{10}(x) = \frac{1}{4,000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n}\cos\left|\frac{x_i}{\sqrt{i}}\right| + 1, \ -600 \le x_i \le 600$

Hartmann Function $f_{11}(x) = -\sum_{i=1}^{4} a_i \exp\left[-\sum_{j=1}^{3} A_{ij}(x_j - P_{ij})^2\right], \ 0 < x_j < 1$

where,

$$a = [1, 1.2, 3, 3.2]^T, A = \begin{bmatrix} 3.0 & 10 & 30 \\ 0.1 & 10 & 35 \\ 3.0 & 10 & 30 \\ 0.1 & 10 & 35 \end{bmatrix}, P = 10^{-4}\begin{bmatrix} 6890 & 1170 & 2673 \\ 4699 & 4387 & 7470 \\ 1091 & 8732 & 5547 \\ 381 & 5743 & 8828 \end{bmatrix}$$

Hump Function $f_{12}(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4$, $-5 \le x_i \le 5$

Levy Function $f_{13}(x) = \sin^2(\pi y_1) + \sum_{i=1}^{n-1} \left[(y_i - 1)^2 (1 + 10 \sin^2(\pi y_i + 1)) \right.$
$\left. + (y_n - 1)^2 (1 + 10 \sin^2(\pi y_n)) \right]$

$$y_i = 1 + \frac{x_i - 1}{4}, \quad -10 \le x_i \le 10$$

Matyas Function $f_{14}(x) = 0.26(x_1^2 + x_2^2) - 0.48x_1x_2$, $-10 \le x_i \le 10$

Michalewics Function $f_{15}(x) = - \sum_{i=1}^{2} \sin(x_i) \sin(ix_i^2/\pi)^{2m}$, $m = 10$, $0 \le x_i \le \pi$

Perm Functions $f_{16}(x) = \sum_{k=1}^{n} \left[\sum_{i=1}^{n} (i^k + \beta) \left((x_i/i)^k - 1 \right) \right]^2$, $\beta = 0.5$, $-n \le x_i \le n$

Powell Function $f_{17}(x) = \sum_{i=1}^{n/4} (x_{4i-3} + 10x_{4i-2})^2 + 5(x_{4i-1} - x_{4i})^2$
$+ (x_{4i-2} - x_{4i-1})^4 + 10(x_{4i-3} - x_{4i})^4$, $-4 \le x_i \le 5$

Power Sum Function $f_{18}(x) = \sum_{k=1}^{n} \left[\left(\sum_{i=1}^{n} x_i^k \right) - b_k \right]^2$, $b = (8, 18, 44, 114)$,
$-4 \le x_i \le 5$

Rastrigin Function $f_{19}(x) = \sum_{i=1}^{n} (x_i^2 - 10\cos(2\pi x_i) + 10)$, $-5.12 \le x_i \le 5.12$

Rosenbrock Function $f_{20}(x) = \sum_{i=1}^{n} \left(100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2 \right)$, $-5 \le x_i \le 10$

Schwefel Function $f_{21}(x) = 418.9829n - \sum_{i=1}^{n} \left(x_i \sin \sqrt{|x_i|} \right)$, $-500 \le x_i \le 500$

Shekel Function $f_{22}(x) = - \sum_{j=1}^{m} \left[\sum_{i=1}^{4} (x_i - C_{ij})^2 + \beta_j \right]^{-1}$, $m = 10$, $0 \le x_i \le 10$

$$\beta = \frac{1}{10}[1, 2, 2, 4, 4, 6, 3, 7, 5, 5]^T, \quad C$$

$$= \begin{bmatrix} 4.0 & 1.0 & 8.0 & 6.0 & 3.0 & 2.0 & 5.0 & 8.0 & 6.0 & 7.0 \\ 4.0 & 1.0 & 8.0 & 6.0 & 7.0 & 9.0 & 5.0 & 1.0 & 2.0 & 3.6 \\ 4.0 & 1.0 & 8.0 & 6.0 & 3.0 & 2.0 & 3.0 & 8.0 & 6.0 & 7.0 \\ 4.0 & 1.0 & 8.0 & 6.0 & 7.0 & 9.0 & 3.0 & 1.0 & 2.0 & 3.6 \end{bmatrix}$$

Shubert Function $f_{23}(x) = \sum_{i=1}^{5} i\cos[(i+1) \cdot x_1 + i] \cdot \sum_{i=1}^{5} i\cos[(i+1) \cdot x_2$
$+ i]$, $-10 \le x_i \le 10$

Sphere Function $f_{24}(x) = \sum_{i=1}^{n} x_i^2$, $-5.12 \le x_i \le 5.12$

Sum Squares Function $f_{25}(x) = \sum\limits_{i=1}^{n} ix_i^2$, $-10 \leq x_i \leq 10$

Trid Function $f_{26}(x) = \sum\limits_{i=1}^{n} (x_i - 1)^2 - \sum\limits_{i=2}^{n} x_i x_{i-1}$, $-n^2 \leq x_i \leq n^2$

Zakharov Function $f_{27}(x) = \sum\limits_{i=1}^{n} x_i^2 + \left(\sum\limits_{i=1}^{n} 0.5ix_i\right)^2 + \left(\sum\limits_{i=1}^{n} 0.5ix_i\right)^4$, $-5 \leq x_i \leq 10$.

3.2 Results and Discussion

Test results are listed in Table 1. We can find that the results of HS are always not as good as that of ABC and STA, except for f_{11}, f_{15} and f_{23}. It seems that HS are capable of solving problems without much interaction between variables, and the solution accuracy and global search ability of HS are also not satisfactory.

Table 1 Results for three algorithms on benchmark instances

Functions	HS mean ± std	ABC mean ± std	STA mean ± std
$f_1(n = 2)$	0.14 ± 0.57-	8.88E−16 ± 0 ≈	8.88E−16 ± 0
$f_2(n = 2)$	0.35 ± 0.53-	3.61E−06 ± 1.38E−5-	4.31E−11 ± 4.91E−11
$f_3(n = 2)$	0.73 ± 0.62-	0 ± 0 ≈	0 ± 0
$f_4(n = 2)$	0.08 ± 0.14-	4.57E−17 ± 4.90E−17+	4.80E−11 ± 3.99E−11
$f_5(n = 2)$	0.39 ± 0.01-	0.39 ± 5.46E−16 ≈	0.39 ± 1.50E−16
$f_6(n = 4)$	7.20 ± 19.98-	0.21 ± 0.14-	0.001 ± 0.002
$f_7(n = 25)$	12.17 ± 5.19-	7.51E−15 ± 2.71E−15+	0.60 ± 0.20
$f_8(n = 2)$	−0.43 ± 0.49-	−0.9057 ± 0.27-	−1.0 ± 1.31E−11
$f_9(n = 2)$	11.48 ± 12.86-	3.002 ± 0.008-	3.00 ± 4.77E−9
$f_{10}(n = 2)$	0.16 ± 0.14-	0 ± 0 ≈	0 ± 0
$f_{11}(n = 3)$	−3.86 ± 2.88E−8 ≈	−3.86 ± 1.82E−15 ≈	−3.86 ± 2.96E−10
$f_{12}(n = 2)$	2.23E−5 ± 8.82E−5-	4.65E−8 ± 0 ≈	4.66E−8 ± 1.13E−10
$f_{13}(n = 30)$	0.90 ± 0.22-	4.98E−16 ± 5.39E−17+	3.84E−11 ± 4.80E−12
$f_{14}(n = 2)$	0.05 ± 0.06-	4.27E−10 ± 1.75E−9-	1.97E−250 ± 0
$f_{15}(n = 2)$	−1.8013 ± 5.44E−5 ≈	−1.8013 ± 6.83E−16 ≈	−1.8013 ± 1.01E−10
$f_{16}(n = 4)$	5.94 ± 9.22-	0.15 ± 0.14-	0.01 ± 0.03
$f_{17}(n = 24)$	10.27 ± 5.54-	1.88E−4 ± 5.94E−5 ≈	1.13E−4 ± 2.36E−5
$f_{18}(n = 4)$	0.29 ± 0.49-	0.02 ± 0.01-	4.33E−4 ± 5.02E−4
$f_{19}(n = 2)$	0.09 ± 0.30-	0 ± 0 ≈	0 ± 0
$f_{20}(n = 2)$	1.02 ± 1.49-	0.01 ± 0.01-	4.38E−8 ± 1.71E−7
$f_{21}(n = 2)$	0.03 ± 0.16-	2.54E−5 ± 0-	2.54E−5 ± 1.48E−12
$f_{22}(n = 4)$	−5.61 ± 3.41-	−10.53 ± 9.34E−5 ≈	−10.53 ± 3.06E−10
$f_{23}(n = 2)$	−186.73 ± 5.09E−4 ≈	−186.73 ± 3.57E−14 ≈	−186.73 ± 3.28E−8
$f_{24}(n = 30)$	0.72 ± 0.20-	5.08E−16 ± 5.69E−17-	0 ± 0
$f_{25}(n = 20)$	0.69 ± 0.50-	2.58E−16 ± 3.72E−17-	0 ± 0
$f_{26}(n = 10)$	−78.37 ± 113.69-	−210 ± 7.32E−7 ≈	−210 ± 1.86E−10
$f_{27}(n = 2)$	6.37E−4 ± 2.80E−3-	2.91E−18 ± 2.55E−18-	0 ± 0
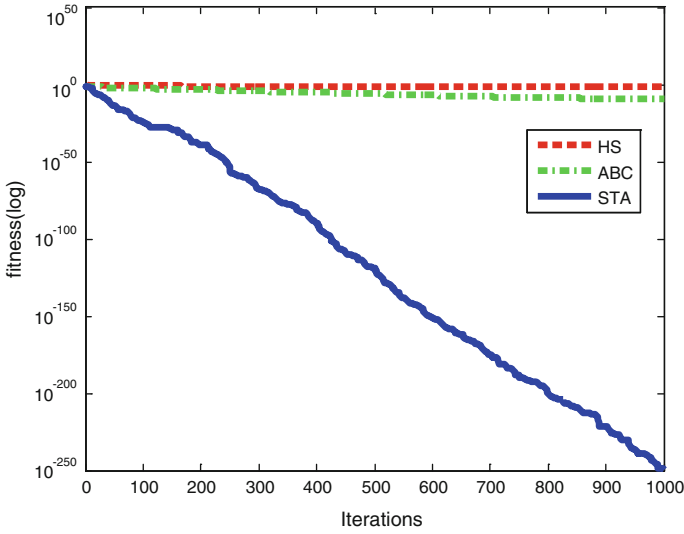
Fig. 1 The average fitness curve of Matyas function by HS, ABC and STA

For ABC and STA, we can find their results are much more satisfactory, and they are able to obtain the global solutions for the majority of the test problems. To be more specific, we can find that ABC outperforms STA for f_4, f_7 and f_{13}, and it can gain higher precision than STA, especially for f_7, which indicates that ABC are more suitable for problems with strongly interacted structure. On the other hand, for f_2, f_6, f_8, f_9, f_{14}, f_{16}, f_{18}, f_{20}, f_{24}, f_{25} and f_{27}, STA outperforms ABC in terms of solution accuracy, which indicates STA has stronger local exploitation ability than that of ABC.

Figure 1 gives the average fitness curve of Matyas function by the three algorithms. We can find that STA is more capable of searching in depth.

4 Conclusion

In this paper, we investigate the mechanisms and performances of state transition algorithm, harmony search and artificial bee colony. Similarities and differences of the algorithms are mainly focused. A suit of unconstrained optimization problems have been used to evaluate these algorithms. Experimental results show that both state transition algorithm and artificial bee colony have better global search capability and can achieve higher solution accuracy than harmony search, artificial bee colony is more capable of solving problems with strongly interacted variables, and state transition algorithm has the potential ability to search in depth.

References

1. Geem ZW, Kim JH, Loganathan GV (2001) A new heuristic optimization algorithm: harmony search. Simulation 76(2):60–68
2. Lee KS, Geem ZW (2005) A new meta-heuristic algorithm for continuous engineering optimization: harmony search theory and practice. Comput Methods Appl Mech Eng 194:3902–3933
3. Yang XS (2009) Harmony search as a metaheuristic algorithm. In: Music-inspired harmony search algorithm: theory and application. Springer, pp 1–14
4. Karaboga D, Basturk B (2007) A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. J Glob Optim 39:459–471
5. Karaboga D, Akay B (2009) A comparative study of artificial bee colony algorithm. Appl Math Comput 214:108–132
6. Zhou XJ, Yang CH, Gui WH (2011) Initial version of state transition algorithm. In: 2nd international conference on digital manufacturing and automation (ICDMA), pp 644–647
7. Zhou XJ, Yang CH, Gui WH (2011) A new transformation into state transition algorithm for finding the global minimum. In: 2nd international conference on intelligent control and information processing (ICICIP), pp 674–678
8. Zhou XJ, Yang CH, Gui WH (2012) State transition algorithm. J Ind Manag Optim 8(4):1039–1056
9. Yang CH, Tang XL, Zhou XJ, Gui WH State transition algorithm for traveling salesman problem. To be published in the 31st Chinese control conference, arXiv: 1206.0329
10. http://www-optima.amp.i.kyotou.ac.jp/member/student/hedar/Hedar files/TestGO files/Page 364
11. https://sites.google.com/a/hydroteq.com/www/
12. http://mf.erciyes.edu.tr/abc/

Data Acquisition and Signal Processing for Endless Rope Continuous Tractor Monitoring

Wei Chen and Hai-shun Deng

Abstract Through datas acquisition, processing the data and signal, to achieve full, real-time monitoring of running state of endless rope continuous tractor. Research and analysis the software and hardware of the endless rope continuous tractor monitoring, it has the function of the winch position and speed of real-time display, leakage communication, language broadcaster, fault logging, winch driver identification, fault masking, two-speed winch speed selection. It can be comprehensive protected for emergency stop to lockout, overwinding protection in the head and tail in running, and remote monitoring through the host computer.

Keywords Monitoring · Endless rope continuous tractor · Data acquisition · Signal processing

1 Introduction

Fully mechanized coal faces with small winch combined transport with the car battery, you need a lot of manpower and material resources, the factor of safety is low, accident rate is high, transport efficiency is low and high operating costs.

Endless rope continuous tractor is one kind of novel emerged in recent years and is quick to promote the coal mine auxiliary transportation equipment, endless rope continuous tractor coal mine at home and abroad, has been widely used.

W. Chen (✉) · H. Deng
School of Mechanical Engineering, Anhui University of Science and Technology, Huainan 232001, China
e-mail: ch163w@163.com

But at this stage the majority of endless rope continuous tractor usually only start and stop control by triple button control the continuous tractor motor switch turns on and off the endless rope continuous tractor ordinary sound and light RBI signal. This operation control systems generally do not have to install communication, protection and comprehensive security control device [1].

Endless rope continuous tractor in running condition monitoring and overspeed protection device following major problems:

(1) Endless rope continuous tractor running status, position, speed and other data can not be displayed in real time, according to the situation can not change, and real-time operation;
(2) Can not be based on changes in working conditions, timely adjustment of operating parameters;
(3) Can not find out the endless rope continuous tractor location and type of failure, and failure and the time to record and save data;
(4) Poor communication effectiveness, without warning device, prone to off road, off the ropes; starting current, unstable operation [2];
(5) Various failures can not be real-time protection for endless rope continuous tractor;
(6) No remote communication function, ground centralized control room endless rope continuous tractor operating conditions can not be observed in real time;

Endless rope continuous tractor monitoring system to solve the above problems, the urgent need to develop, to achieve full, real-time monitoring of continuous tractor operation status, and a variety of comprehensive protection its running process.

2 Overall Structure

Endless rope continuous tractor monitoring the overall design of the comprehensive security system will be PLC programmable logic controller for the control of the core, the choice of $10.4''$ color touch screen control unit, by collecting various sensors, wireless the leakage communications signal, the preparation of PLC and touch screen program, to monitor comprehensive security features to meet explosion-proof requirements will be achieved through the operation platform button, various parameter settings and switch the screen [3].

Using the configuration software design, computer monitoring program, comprehensive security system monitoring endless rope continuous tractor through the communication interface is connected to the host computer, to achieve the purpose of remote monitoring.

Increase the winch driver recognition function, increase design fault occurs promptly determine and quick way to find and troubleshooting [4].

Protection device consists of the following five components:

1. Mine endless rope winch integrated protection device host: small integrated programmable controller to control the core, together with the 10.4″ color LCD screen;
2. Mining intrinsically safe speed sensor: Hall switch mode fixed sheave on the winch pressure device rope near the mouth of the operation of the winch speed state passed to the host in the form of a square wave [5].
3. Mining intrinsically safe overwinding switch: the winch head and tail of the installation of the to prevent tractor overtravel happened volume; installed in another closer location away from the nose, as a position display calibration origin switch [6];
4. Mine flameproof and intrinsically safe type voice light alarm device (II): winch along layout, play a role in the emergency stop latching winch, at the same time play the point, talk, radio, voice alarm, general winch along every 100 m to install one;
5. Mine explosion-proof and intrinsically safe voice lights are driving, are not allowed to pedestrian warning is issued when the alarm device (III): installed at the fork in the road, the corners on the downhill slope larger location, tractor after remind pedestrians to take necessary precautions [7].

3 Hardware Part of Data Acquisition

The block diagram of the device shown in Fig. 1.

PLC control box, the device, signal acquisition and processing, and to start and stop the tractor forward and reverse control and protection; to using monitor real-time surveillance tractor operating information; fails, the console will sound and light alarm.

Speed sensor pulse generated by the induction magnet, pulse count complete tractor speed, position monitoring, as well as ultra-under speed the soft limit protection of the nose and tail; nose, tail and origin of each placed a volume switch, throughout by monitoring the volume switch the state to achieve nose the hard tail overwinding protection and location of cumulative error calibration; every 200 m along the arrangement of a signal emergency stop control box, to be completed by the state of the emergency stop button along the emergency stop lockout function [8];

Voice light alarm device installed at the fork in the road, the corners, on the downhill slope larger, etc. tractor run will advance 50 m issue is driving, not allowed to pedestrian alarm to warn pedestrians to take necessary precautions. The hardware part of data acquisition as show in Fig. 2.
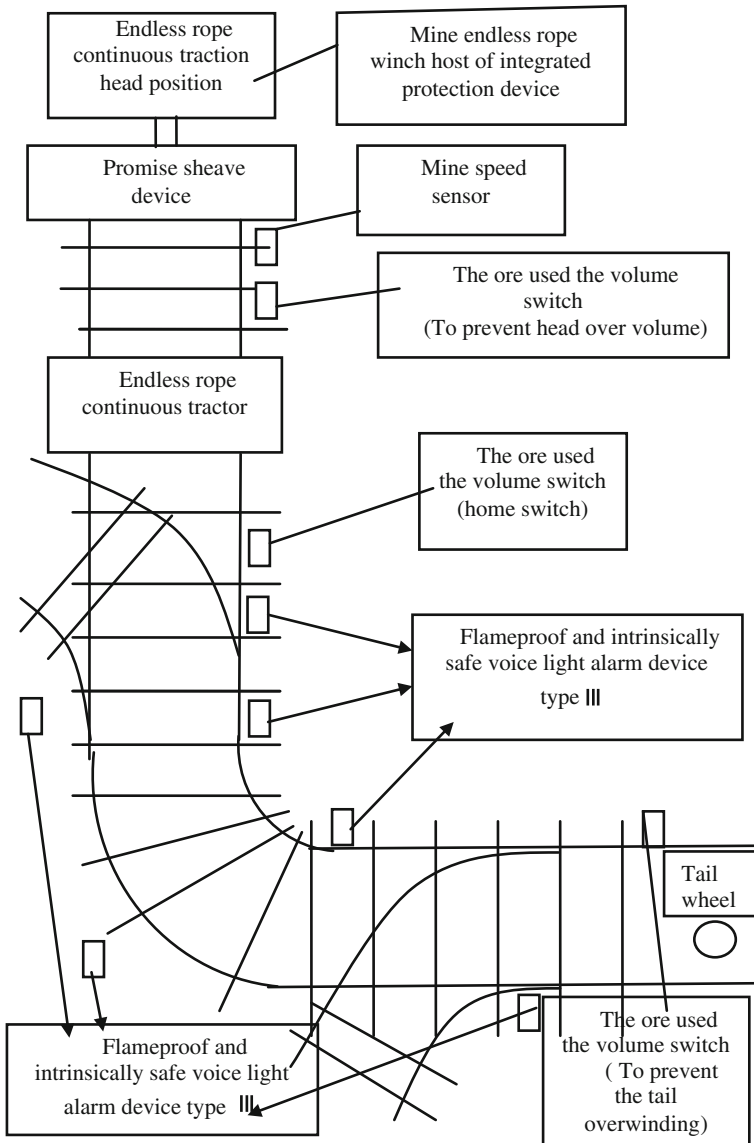
Fig. 1 Unit block diagram

4 Data and Signal Processing

Protection device host has a five-button, followed by "running forward", "reverse run" and "stop", "forward jog and reverse jog", play a positive, reverse start-stop or reverse jog winch role [9].
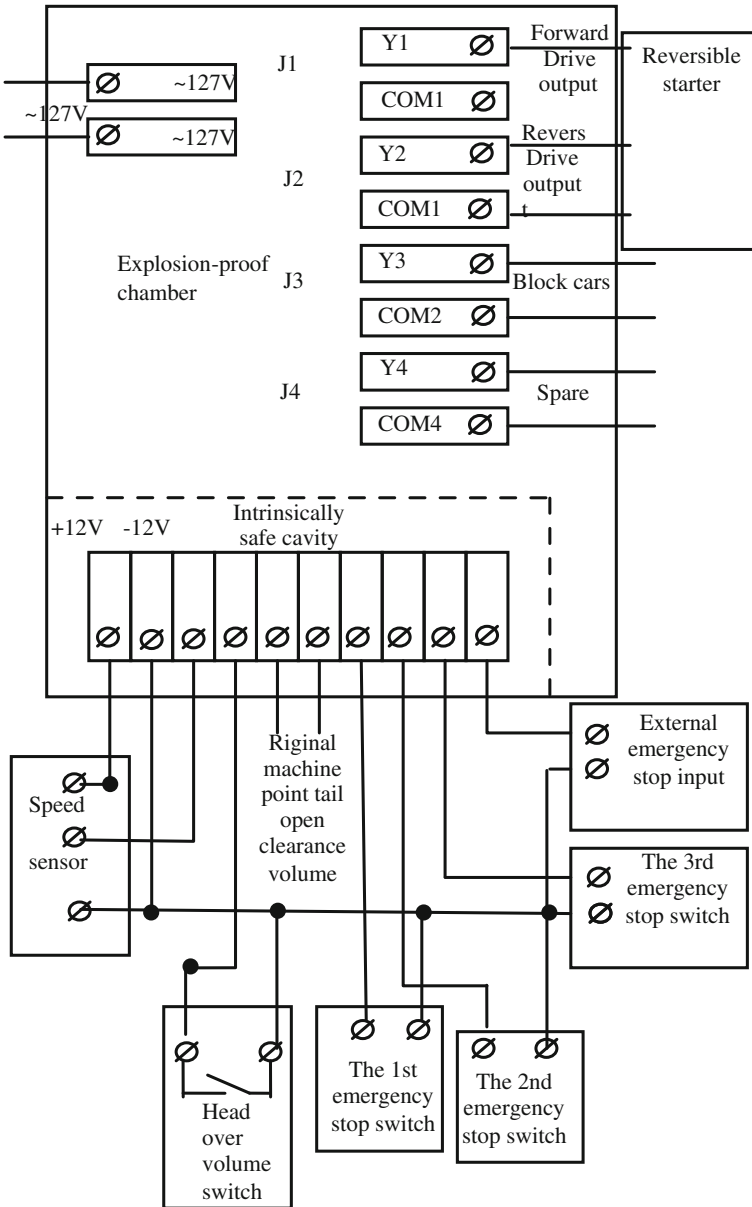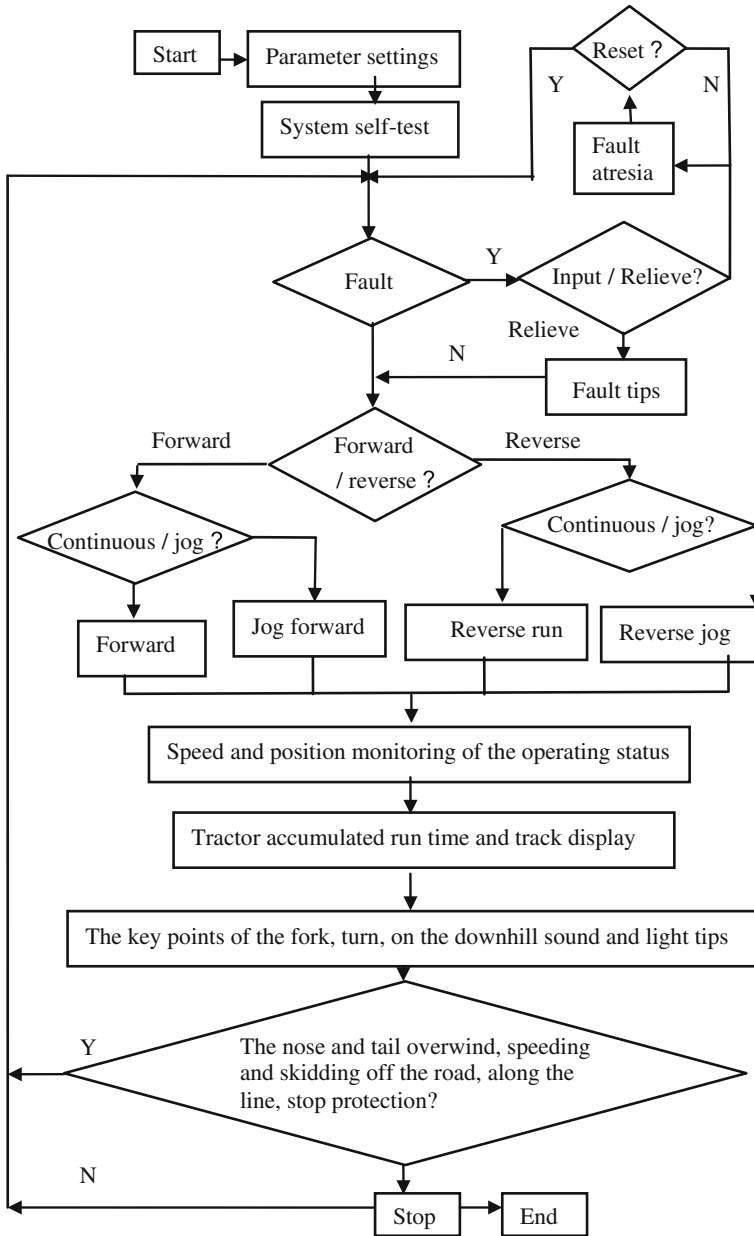
Fig. 2 Hardware part of data acquisition

Fig. 3 Process flow diagram

Data and work processes based on the collected data and process requirements, through the PLC programming, program flow chart below shows in Fig. 3.

5 Conclusion

Device through the wireless transmission speed sensor, data acquisition, storage, analysis, processing, display and wireless communication, the correct judgment to achieve effective monitoring of endless rope continuous tractor running state and overspeed fault handling. Endless rope winch monitoring system.

Winch driver can place the control room, real-time view of the health of the shuttle car. (1) to improve the production conditions of the employees of the coal mine. Through remote monitoring to facilitate security inspection staff to monitor the production of endless rope winch, through the collection of field information to improve the efficiency of communication between each other.

References

1. Rzesiak LM, Kaszewski A (2006) Canbus commumcations m Hybrid power generator HYGEN. Przegl Elektrotechniczny 82(2):128–132
2. Pirlll0 LM, Vasque F (2003) Reliable real-time communication in CAN networks. IEEE Trans Comput 52(12):l594–1607
3. Vladimir O, Nikiforo V (2001) Nominear servocompensation of unkown extemaldistul. Bances Automatica 2001:1647–1653
4. Yang H (2008) The JWB mine endless rope winch usage. Mech Manage Dev 4:62–63
5. Cao R, Strong OZ, Duzhong Q (2010) The comprehensive protection device in the ZBJW endless rope winch. Coal 3:58–59
6. Xu G, Jiang J, Zhang W (2009) The endless rope winch and leaks communication device in the application of mechanized mining face. Coal 3:61–64
7. Chen H, Ma XP, Wang C, Yu H, Zou X (2005) Endless rope continuous tractor operation status monitoring and over speed protection system management and maintenance of equipment 19:48–50
8. Zhang HP (2012) Wireless video surveillance the endless rope winch system. Energ Technol Manage 1:138–140
9. Ahmadi I (2011) Dynamics of tractor lateral overturn on slopes under the influence of position disturbances. J Terrramech 48(5):339–346

A Subcarriers and Bits Allocation Scheme for Multi-User OFDMA System

Jingxue Ran and Bo Xiao

Abstract In Multiple-In Multiple-Out (MIMO) Broadcast systems with limited feedback, when receivers are equipped with multiple antennas, the so-called quantization error of channel direction matrix were unavoidably introduced. This quantization error will result in the reduction of system capacity since it was referenced in the beam forming design. In this paper, we proposed a robust beam forming method based on MMSE, which considers the statistical characteristics of the quantization error for channel direction matrix. Simulation results show that the proposed method effectively enhanced system capacity, reduced bit error rate (BER), as well as mitigated the Plateau Effect.

Keywords OFDMA · Bits allocation · MIMO

1 Introduction

With the development of wireless networks, the services of wireless networks convert from voice service to data service especially multimedia services. How to satisfy the demand of different users and different styles has been a question which is in dire need of solution. Orthogonal Frequency Division Multiplexing (OFDM) has been a core technology in the next generation wireless networks, due to the capability of keeping out interference, high bandwidth utilization and converting the selective fading channel to several flat fading sub channels.

J. Ran (✉) · B. Xiao
Modern Education Technical Department, Minzu University of China,
Beijing 100081, China
e-mail: ranjingxue@163.com

B. Xiao
e-mail: zhiqing7057899@163.com

OFDMA is a kind of multiple access technology based on OFDM. It can improve the channel capacity through designing the proper sub carriers and bits allocation algorithm. In OFDM system, single user occupies all of sub carriers. However, in OFDMA system, multi-users occupy all of sub carriers. These will be the question that how to allocate the resources to users, especially the sub carriers allocation.

In multi-user OFDMA system, due to the location of multi-users are independent, the users' frequency selective fading channels are also independent. Hence, the sub carriers are different among the users. And the users' channels are different at the same sub carrier. Some of users' channels are in deep fading, and some of users' channels are in low fading. If the sub carriers are allocated to the low fading all the time, the every carrier can be used efficiently, so the system capacity is higher. Meanwhile, allocating the more power and bits to the low fading users and the less power and bits to the deep fading users will be improve the system's reliability and efficiency.

In multi-user OFDMA system, under satisfying the users' demand, minimizing the transmitting power, named Margin Adaptive (MA) problem and several subcarriers may be good for users. However, those subcarriers are not shared simultaneously. Due to the mathematics models to allocate the subcarriers and bits are not convex; so many researches are mainly on designing the suboptimal algorithms to obtain the better results. For examples, fixed bits and subcarriers allocation designing schemes, two-step bits and subcarriers allocation designing schemes etc. These algorithms which decompose the joint allocation schemes to independent steps schemes make the problem simple, but the performances loss will be large. Wong et al. [1] improve the integer bits allocation schemes to the real number bits allocation schemes, this scheme utilizes Lagrange formula and greedy algorithm to allocate the subcarriers and bits. However, this scheme complexity is high. The algorithms of [2] and [3] are related with the order of allocation.

2 System Model

Multi-user OFDM downlink system depict as Fig. 1, including K users and N subcarriers [4]. Assuming the bandwidth of subcarriers is much smaller than coherent bandwidth of channels, so the pulse responses of channels are flat. Users obtain the channels' gain through channel estimation and used for the channel state information [5] (CSI), and feed back to base station through special channels. The station based on the CSI allocates the subcarriers with different users, and decide the power and bits of each subcarrier.

Assuming the BER of the user $k(1 \leq k \leq K)$ is B_k, the data rate is R_k, the power of subcarrier n is $P_{k,n}$, the number of bits are $r_{k,n}(r_{k,n} \in \{r\})$, then,
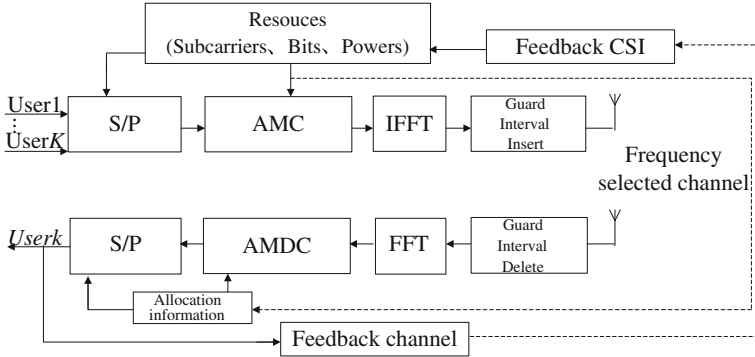
Fig. 1 Multi-user OFDM downlink system model

$$r_{k,n} = \log_2\left(1 + P_{k,n}g_{k,n}\right)$$
$$R_k = \sum_{n=1}^{N} r_{k,n} \quad (k = 1\dots K) \tag{1}$$

In above formula $g_{k,n} = \frac{h_{k,n}^2}{\Gamma_k N_0}$ ($g_{k,n}$ is the equivalent channel gain), $h_{k,n}^2$ is the channel gain in subcarrier n of user k. In order to avoid interferences, each user must use the subcarrier only. Using mathematical model is

$$If\, r_{k,n} \neq 0 \; then \; r_{k',n} = 0, k \neq k' \tag{2}$$

The MA problem is constraint under $\{R_k, BER_k\}$, based on (1) and (2), minimizing the total transmitting power, then

$$P_{\min} = Min \sum_{k=1}^{K}\sum_{n=1}^{N} P_{k,n} = \sum_{k=1}^{K} P_k^{\min} \tag{3}$$

From (3), we can know that minimizing $P_{\min}$ is equal to minimize the power $P_k^{\min}$ of each user. Assuming C_k is the set of subcarriers for user k. $|C_k|$ is the quantity of the set C_k. Based on water filling theory [6], the power of user k for subcarrier n

$$P_{k,n\in C_k} = \left(\frac{1}{\lambda} - \frac{1}{g_{k,n\in C_k}}\right)^+ \tag{4}$$

Then

$$\frac{1}{\lambda} = \frac{1}{|C'_k|}\left(P_k + \sum_{n\in C'_k} \frac{1}{g_{k,n}}\right) \tag{5}$$

where $C_k' \subset C_k$ and $P_{k,n \in C'} \neq 0$, C_k' is the subcarrier set of the user k after water filling. Now, (1) (4) (5) lead to $(C_k', R_k) \rightarrow (P_k^{\min}, \mathbf{P}_k, \mathbf{r}_k)$, where $\mathbf{P}_k = \{p_{k,n \in C_k}\}$, $\mathbf{r}_k = \{r_{k,n \in C_k}\}$.

Then the minimizing power of the rate for user k is

$$P_k^{\min} = \left| C_k' \right| \left(\frac{2^{R_k}}{\prod\limits_{n \in C_k'} g_{k,n}} \right)^{\frac{1}{|C_k'|}} - \sum_{n \in C_k'} \frac{1}{g_{k,n}} \tag{6}$$

From (6), we can know for any subcarrier $n \in C_k'$, demanding

$$\prod\nolimits_{n \in C_k'} g_{k,n} \Big/ 2^{R_k} < g_{k,n \in C'}^{\ |C_k'|} \tag{7}$$

When user k need add a subcarrier n^*, water filling again, water filling factor is $1/\lambda^*$, power is $P_k^{\min *}$.

3 The Proposed Scheme of Bits and Subcarriers' Allocation

In Seong et al. [7], the scheme proposed power and bits allocation, which based on minimizing transmit power criterion. This algorithm is made to convergence. It allocated the power and bits in two steps. First, salute the iterative formula of the water filling without prefixed step length and initial values which makes the transmitted speed vibrates around prefixed initial value. Then, allocate the final power and bits, based on twice water filling iterative bits allocation, with simple greedy algorithm.

In Li et al. [8], for each allocation, compare each user power, allocate the max channel and gain to the user for max power. In general, allocate the subcarriers one by one to decrease $P_k^{\min}$, with minimizing the total transmitted power $P_{\min}$. Finally, it is achievable to minimize the total transmit power. The detail steps as follows:

Initialization, subcarriers set is $\mathbf{N} = \{n|n = 1 \ldots N\}$, the number of subcarriers is $|\mathbf{N}| = N$, the user set is $\mathbf{U} = \{k|k = 1 \ldots K\}$, the number of users is $|\mathbf{U}| = K$; the subcarrier set for each user is $C_k = \Phi$, and $|C_k| = 0$. In order to guarantee transmitting demand, the minimizing number of the subcarriers is $|C|_{k,\min} = R_k/\log_2 M$ and the maximizing number of the subcarriers is $|C|_{k,\max} = R_k/2$.

Allocate the largest channel gain subcarrier to each user, assuming the subcarrier rate is $R_k(R_k > \log_2 M)$.

In purpose of minimizing transmit power; allocate the subcarriers one by one. For each time, allocate the max channel gain subcarrier n^* to the user k^*, $k^* \in \mathbf{U}$ of max power. Then water filling again, from (1)–(5), we obtain the $P_{k^*}^{\min}$.

Judging residual subcarrier number is zero or not. If $N \neq \Phi$, repeat step 3 until $N = \Phi$.

Make greedy algorithm to revise bits. Because of power water filling obtain the number of bit is continuous, at integer $r_{k,n}$, $r_{k,n} \in \{0, 2, \ldots, \log_2 M\}$, the residual bits is $R_k^L = R_k - \sum_{n \in C_k} r_{k,n}$ in C_k with greedy algorithm, end.

4 Simulation Results and Analysis

In all simulations presented in this section, we simulate the system capacity in 8 K voice service. Reference to PHY simulation result, the purpose SINR set the BER as 10^{-3}. In system simulation, we consider path loss, shadow fading and fast fading, no considering the influence of handoff. In order to simulate the dynamic resource allocation, assuming the users achieve in average, end to the number of user achieve the predefined number. If the SINR is lower than the target value 0.5 dB, we deem the user is satisfied. We simulate in multi-cell macro cellular, The target value of SINR achieve 8 dB, each user need 2 uplink subcarrier and 2 downlink subcarrier. MT maximal transmitting power is 21 dBm; AP maximal transmitting power is 20 W. The system bandwidth is 2×10^6 Hz, cellular radius is 500 m, propagation model is 5.3 GHz channel model, shadow fading is lognormal distribution, and variance is 8 dB.

In Fig. 2, each user transmit 24 bits, BER is 10^{-3}. In Fig. 3, the proposed scheme is much lower than ABMU for total SNR. When 28 users are considered, about 5 dB is lower than ABMU. Compared to Fig. 2, multimedia services are better.
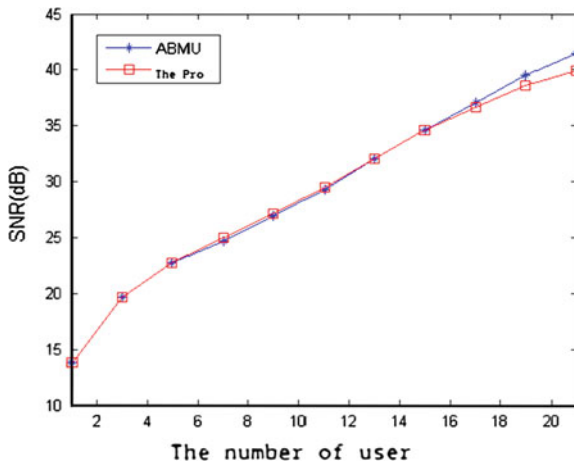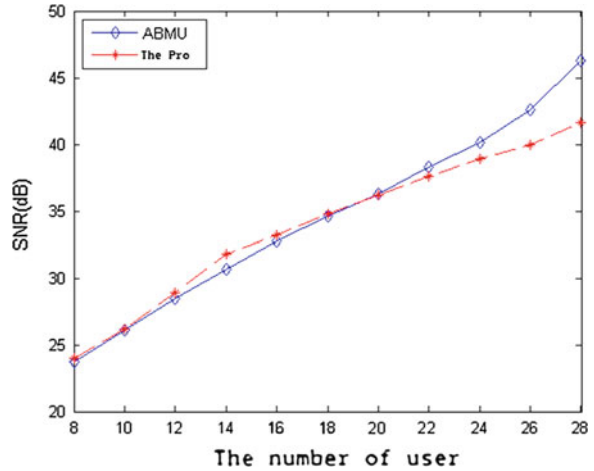

Fig. 2 Total SNR for different user numbers ($R_k = 24, BER_k = 10^{-3}, \forall k$)

Fig. 3 Total SNR for
different user numbers
($R^1 = 12$, $R^2 = 24$,
$BER^1 = 10^{-3}$, $BER^2 = 10^{-4}$)

From the simulation result, the scheme is proper to the rate higher, user more system. From Figs. 2 and 3, the proposed scheme is better than ABMU, because the proposed scheme considers the co-subcarrier interference, chooses the minimal interference subcarrier and user, decreases the transmitting power, improves the system capacity and satisfies the users' QoS.

5 Conclusion

In this paper we have proposed schemes for constraint interference dynamic subcarrier allocation. The scheme is suitable for multiuser wireless system. We have shown that the scheme combines with power control. The scheme decreases the transmitting power we also have shown the proposed scheme improves the system spectrum efficiency and system user satisfaction degree comparison.

References

1. Wong CY, Cheng RS, Letaief KB et al (1999) Multiuser OFDM with adaptive subcarrier bit and power allocation. IEEE J Sel Areas Commun 17(10):1747–1758
2. Zhang G (2004) Subcarrier and bit allocation for real-time services in multiuser OFDM systems. IEEE ICC 5:2985–2989
3. Yu G, Zhang Z, Chen Y (2006) A novel resource allocation algorithm for real-time Services in multiuser OFDM systems. IEEE VTC 3:1156–1160
4. Rye S, Ryu BH, Seo H, Shin M, Park S (2005) Wireless packet scheduling algorithm for OFDMA system based on time-ultility and channel state. ETRI J 27(6):344–356
5. Wang Y, Tian W, Zhang P (2004) Two-dimensional resource allocation for OFDM/TDMA microcellular networks. IEEE VTC, 2004-fall, 5:3452–3456

6. Parag P, Bhashyam S, Aravind R (2000) A subcarrier allocation algorithm for OFDMA using buffer and channel State information. In: Proceedings of IEEE vehicular technology conference, VTC2005-fall, Dallas
7. Seong K, Mohseni M, Cioffi JM (2006) Optimal resource allocation for OFDMA downlink systems. In: Proceedings of ISIT 2006, Seattle, USA, 1394–1398 July 2006
8. Li G, Liu H (2003) Downlink dynamic resource allocation for multicell OFDMA system, Vehicular technology conference, 2003. VTC 2003-Fall, vol 3, pp 1689–1702, 6–9 Oct 2003

Study on the Prediction of Gas Content Based on Grey Relational Analysis and BP Neural Network

Jiajia Lu, Ping Chen, Jinshan Shen, Zepeng Liang and Huaze Yang

Abstract The paper mainly illustrated a kind of BP Neural Network model in MATLAB Neural Network Toolbox to predict gas content of the coal seam based on analyzing grey relational degree. This model constructed a method of gas content prediction by choosing four dominate effect factors (Coal seam buried depth, Geologic structure, Roof lithology, Coal seam thickness) as the input parameters. It has been established for training and testing, and forecasting the gas content of coal seam by using the learning samples which were collected from the instances of typical exploited borehole data of Panyi East coal mine in Huainan coal mining area. The results show that the model is an efficient prediction method for gas content, and its prediction accuracy and feasibility are better than the traditional predicting methods. It also can better meet the practical requirements of safety production in coal mine including provide some references for mine gas disaster prevention.

Keywords Grey relational analysis · BP neural network · Effect factors · Forecasting · Gas content

1 Introduction

Coal and gas outburst is one of the important disasters threaten mine production and safety [1]. Gas content in coal seam is one of the important basis parameters with predicting coal and gas outburst. Also it is an important parameter which is used to predict and prevent gas disaster. Therefore, the study on predicting the gas content in the coal seam would not only instruct to apply outburst prevention

J. Lu (✉) · P. Chen · J. Shen · Z. Liang · H. Yang
School of Earth and Environment, Anhui University of Science and Technology,
Huainan 232001, China
e-mail: jiajia86217@163.com

measures scientifically and reduce its related work amounts, but also guarantee the personal safety of the outburst layer owing to the continuous checking on the outburst risk of the outburst layer [2]. So the forecast of gas content has great practical significance.Gas content is controlled by many geological factors, and it has a complex non-linear relationship with various effect factors. In the actual mining production, gas has obvious grey information owing to influence of local geological factors [3]. This paper used Grey Relational Analysis to determine optimal influence factors of gas content and lay the foundation for confirming the input layer through Neural Network. It took No.13-1 coal bed of Panyi East mine for instance, combined Grey Relational Analysis with BP Neural Network to built gas content prediction model in order to achieve forecasting the gas content of without mining area.

2 Analysis of Effect Factors

The kinds of factors affecting coal seam gas deposit and distribution are mainly as follows: coal seam buried depth, geologic structure, roof lithology, coal seam thickness, volatile component, etc. Before the study of the relationship between various effect factors and gas occurrence, we had to translate the qualitative description factors into quantitative factors based on quantifying its nature and influence of gas occurrence scientifically [4].

1. Coal seam buried depth

In general, the deeper coal seam is, the bigger vertical pressure will be. The bigger vertical pressure gets, the harder gas dissipation will be. The huge depth makes the gas be protected well, and have better gas occurrence.

2. Geologic structure

The faults in geologic structure have destroyed the continuity of coal seam creating the conditions for gas dissipation. Tension-shear faults are conducive to gas dissipation; compression-shear faults go against gas dissipation. Closed anticline which is a good gas tank is beneficial to gas deposit. Panyi East coal field research area is located in the fault of footwall F5, hanging wall F2. There are 35 faults in No.13-1 coal bed according to borehole and three-dimensional seismic. Among them, 22 faults are normal faults, 13 faults are reverse faults. Through choosing the nearest horizontal distance of faults from borehole to coal points as formation parameters when quantifying the faults in view of influence from fault to gas occurrence.

3. Roof lithology

The gas whose dissipation has a very close relationship with the coal seam surrounding rock has different dissipation degrees from a long span of geological time. When the roof lithology is integrated and dense, the gas in coal seam can be

Table 1 Original data of gas content and each influencing factor

Sample number	Gas content (m³/t)	Buried depth (km)	Geologic structure (km)	Roof lithology	Coal seam thickness (m)	Volatile component (%)
1	21.92	683.32	2.52	3	4.54	41.08
2	5.34	770.24	0	2	2.55	46.48
3	19.33	671.82	1.57	3	4.73	39.59
4	21.37	703.84	1.79	3	4.89	43.27
5	19.28	661.30	0.49	1	4.78	41.77
6	16.46	619.90	0.12	3	4.04	42.87
7	18.58	704.24	1.39	3	5.93	30.14
8	12.34	695.80	0.59	1	3.52	21.17
9	11.52	768.51	0.12	3	3.46	31.22
10	3.53	758.85	0.18	1	3.20	41.42
…	…	…	…	…	…	…
25	8.96	730.12	0.35	1	5.23	41.56
26	14.82	806.81	0.32	3	5.09	30.66
27	14.58	652.01	0.15	1	4.52	41.77
28	13.42	743.55	1.68	2	6.93	32.42
29	19.58	636.71	0.35	3	4.69	40.78
30	6.08	745.66	0.79	2	3.45	42.12

preserved easily. Rather, it can escape easily when the roof lithology is porous or fissured. The roof lithologic characters of mine area mainly are siltstone, sandy mudstone and mudstone. In the following Table 1 they are replaced by 1, 2, and 3, respectively.

4. Coal seam thickness

The Coal seam thickness is one key factor that affects gas content. Generally speaking, gas generation will get bigger with the increase of the coal seam thickness. No.13-1 coal seam thickness range is between 2.55 and 6.93 m whose average is 4.48 m, having major gas generation.

5. Volatile component

Coal is a natural adsorption body. Generally, the lower Volatile component is, the higher metamorphic grade is, and the higher its coal rank is, the stronger its gas storage capacity is in ideal environment.

6. Other geologic effect factors

The influence of gas content is faint because the change of coal quality in study area is not obvious and considering the monotony of the hydrogeological condition, these factors can be ignored during model build.

Based on the above factors, this paper collected some typical borehole data from Panyi East coal mine as the samples set of grey relation analysis model. Values of various factors are shown in Table 1.

3 Grey Relational Analysis

The basic idea of grey correlation model is to determine the similar degree of the two sequences based on the sequence curve geometry, the closer the curve shows that the correlation degree of the corresponding sequences is relatively larger, and the smaller the other hand. It can be seen that, to some extent, Grey Relational Analysis can make up the shortcoming of the traditional mathematical methods of statistical analysis. Therefore, it is significant to use the grey correlation method to analyze the relationship between the gas content and the effect factors.

This paper made the factor that is gas content as the reference sequence (controller sequence) $X_0(k)\, k = 1, 2, \cdots, 30$, and other various factors as compared sequence (sub sequence) $X_i(k)\, i = 1, 2, \cdots, 5$.

At first, construct an initial decision matrix. Determining that there are 6 data sequences characterized by 30 criteria:

$$(x_0, x_1, \ldots x_5) = \begin{bmatrix} x_0(1) & x_1(1) & \ldots & x_5(1) \\ x_0(2) & x_1(2) & \cdots & x_5(2) \\ \vdots & \vdots & & \vdots \\ x_0(30) & x_1(30) & \cdots & x_5(30) \end{bmatrix} 30 \times 6 \tag{1}$$

Because of different physical meanings of each factor, the data dimensions have different evaluation criterion. So date normalization should be done before doing Grey Relational Analysis. Normalization equation is:

$$x_i(k) = \frac{X_i(k) - X_{\min}(k)}{X_{\max}(k) - X_{\min}(k)} \tag{2}$$

Calculate the grey relational coefficient for each entity:

$$\xi_i(k) = \frac{\min_i \min_k \Delta_i(k) + \rho \max_i \max_k \Delta_i(k)}{\Delta_i(k) + \rho \max_i \max_k \Delta_i(k)} \tag{3}$$

Among them, $\Delta_i(k) = |x_0(k) - x_i(k)|$ is the absolute difference of index k, the distinguishing coefficient ρ in most situations takes the value of 0.5 because this value usually offers moderate distinguishing effects and good stability [5].

A grey relational degree is a weighted sum of the grey relational coefficients given by the following equation:

$$r_i = \frac{1}{n} \sum_{i=1}^{n} \xi_i(k) \quad (k = 1, 2, \cdots, 30) \tag{4}$$

The results are shown in Table 2.

The relating sequence of various factors for influence of gas content shows the affecting degree on the gas content. According to the calculation results of correlation degree above, correlation degree that various factors for gas content orders

Table 2 Relational degree taxis of influencing factors

Effect factors	Correlation degree (r_i)	Ordering
Coal seam buried depth	0.6915	1
Geologic structure	0.6350	4
Roof lithology	0.6654	2
Coal seam thickness	0.6624	3
Volatile component	0.5653	5

are as follows: Coal seam buried depth > roof lithology > coal seam thickness > geologic structure > volatile component.

When $r_i \geq 0.8$, the correlation degree is better; when $0.8 > r_i > 0.6$, the correlation degree is good; when $r_i < 0.5$, there is no correlation [6]. So, we chose four dominate effect factors (Coal seam buried depth, Geologic structure, Roof lithology, Coal seam thickness) whose correlation degree is good as the input parameters of BP Neural Network.

4 Basic Theory of BP Neural Network

4.1 BP Neural Network

Back Propagation Neural Network is a multilayer feedforward networks using error reverse propagation algorithm and it is one of the most widely used Neural Network models. It can learn and store massive input–output model mapping without disclosure and description of mathematical equation on the mapping relations beforehand. The structure of BPNN is shown in Fig. 1.

Fig. 1 The structure chart of BP neural network model
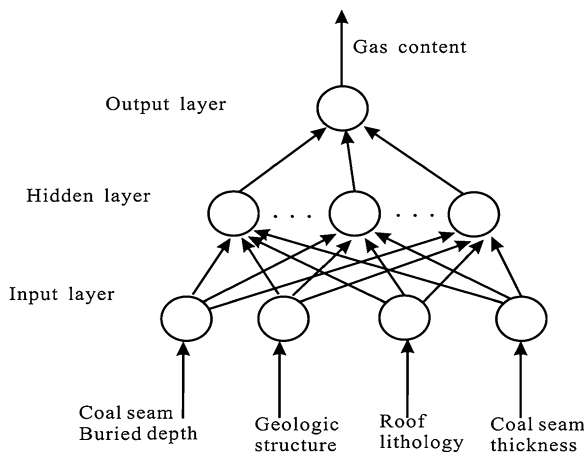
4.2 BP Algorithm of Learning Phase

Set the network input is X, the number of input nodes is n, the number of the hidden layer nodes is p, the relevant transfer function is f_1, the number of the output layer nodes is q, the corresponding transfer function is f_2, the output is Y and the target is O.

1. The forward propagation of information

(a) The output of the jth node in the hidden layer:

$$b_j = f_1 \left[\sum_{i=1}^{n} w_{ji} \cdot x_i + \theta_j \right] \tag{5}$$

(b) The output of the kth neuron in the output layer:

$$y_k = f_2 \left[\sum_{j=1}^{p} w_{kj} \cdot b_j + \theta_k \right] \tag{6}$$

(c) The error function is defined as:

$$e = \frac{1}{2} \sum_{k=1}^{q} (o_k - y_k)^2 \tag{7}$$

2. The backward propagation of error and weight adjustment

(a) Calibration error of output layer node:

$$d_k = (o_k - y_k) \cdot y_k \cdot (1 - y_k) \tag{8}$$

Its correcting value of threshold value is:

$$\Delta \theta_k = \alpha \cdot d_k \tag{9}$$

The weight adjustment of output layer is:

$$w_{kj}(new) = w_{kj}(old) + \Delta \theta_k \cdot b_j \tag{10}$$

(b) Calibration error of hidden layer node:

$$e_j = \left(\sum_{k=1}^{q} w_{kj} \cdot d_k \right) b_j \cdot (1 - b_j) \tag{11}$$

Its correcting value of threshold value is:

$$\Delta \theta_j = \beta \cdot e_j \tag{12}$$

The weight adjustment of hidden layer is:

$$w_{kj}(new) = w_{kj}(old) + \Delta\theta_j \cdot x_i \tag{13}$$

3. Repeat (1), (2) for circulating memory study;

4. According to discriminate the learning results, if the output error is so small that it could be allowed or its maximum cycle times turned up, the learning progress can be terminated.

According to the theory of BP Neural Network and the characteristics of core data, prediction of coal seam gas content can be divided into five steps as follows:
(1) Select the feature vectors of input and output; (2) Select the learning samples; (3) Preprocess the core data; (4) Input sample data to BP network, train and test the BP network; (5) Input the sample data to predict the coal seam gas content.

5 Prediction of Gas Content in Coal Seam

5.1 The Design of Prediction Model

According the above analyses, this paper chose four dominant effect factors from Gray Relational Analysis: coal seam buried depth, geologic structure, roof lithology and coal seam thickness as the input nodes.

The number of the hidden layer nodes is very important. If the number of the nodes is too less, the network can't establish complicated judgment boundary; If the number of the nodes is overabundant, the network will lose the summarizing and judging ability [7]. The number of 6, 9 and 13 are tried in this paper and finally the best number of hidden layer is 9 whose network performance is better.

The gas content was used to be the output node. Therefore, we used 4–9–1 three layers BP network to create a prediction model of coal seam gas content which can be shown in Fig. 1.

5.2 The Preparation of Learning Sample Data

The instances of typical exploited borehole data of Panyi East coal mine must be chosen as sample and their data must be standardized before network training. The learning samples are shown in Table 1.

The parameter values of Input nodes are different and the dimensions diverge greatly, so the values need to be normalized in order to prevent the information of small values from being weakened by big ones. Because the function sigmoid curve changes is smooth between 0 and 0.1 or between 0.1 and 0.9. So the good normalized

Table 3 Contrast between neural network and multiple linear regression models

Sample number	Gas measured value (m^3/t)	Forecasting value by BPNN (m^3/t)	Relative error of BPNN (%)	Multiple regression value	Relative error of multiple regression (%)
8	12.34	12.386	0.373	10.425	12.34
9	11.52	11.530	0.087	10.128	11.52
26	14.82	14.826	0.040	10.685	14.82
28	13.42	13.408	−0.089	16.957	13.42
29	19.58	19.544	−0.184	19.958	19.58
30	6.08	6.166	1.414	10.158	6.08

value range should be [0.10, 0.90]. The formula of $(0.9 - 0.1)\left(\frac{X_i - X_{\min}}{X_{\max} - X_{\min}}\right) + 0.1$ can satisfy the normalized requirements. Quantificational data can be normalized using the above method [8].

5.3 Sample Training and Testing

We chose tansig, logsig and purelin from Neural Network Toolbox in the software MATLAB separately as the transfer function of input layer, hidden layer and output layer. Due to its higher convergence speed and smaller network training error, trainlm function is suggested to be used in training. The biggest times of training are 15,000, the expected error is 0.0001 and learning rate is 0.1.

In 30 sets of data, 24 sets of data as training samples, 6 sets of data (8, 9, 26, 28, 29 and 30) as the testing sample from Table 1. After training, we compared the forecasting results with that based on the traditional gas content forecasting model using multiple regressions. The results were shown in Table 3.

As can be seen from Table 3, the fluctuation of the relative error of BPNN is smaller, which implies that the precision and feasibility of the BPNN forecasting are higher than the traditional multiple regression model. These results indicate that the BPNN model is appropriate for forecasting of the coal seam gas content.

6 Conclusions

1. The effect factors of gas content are so many that five main effect factors were analyzed in this paper, but four dominate effect factors were collected according to the Grey Relational Analysis and we obtained the correlation order of effect factors affecting gas content.
2. The combination of grey theory and neural network has better forecasting precision, which can describe the internal regularity between gas content and other effect factors greatly and predict the coal seam gas content better.

3. There are complex nonlinear relationships between the main controlling factors and the gas productivity, which is difficult to deal with by means of traditional mathematical methods. Artificial neural network has strong nonlinear approximation ability, the accuracy and feasibility of the BPNN forecasting are higher than the traditional multiple regression model and the BPNN model is appropriate for the analysis and forecasting of the coal seam gas content.

References

1. Li C, He X (2005) Prediction method of coal and gas outburst dangerous level in coal roadway face. J China Univ Min Technol 34(1):74–76
2. Nie B, Dai L, Yan A et al (2010) The Study on prediction of coal seam gas content based on support vector regression. China Safety Sci J 6:28–32
3. Ye Q, Lin B (2006) Applicatipon of Gery Theory in predicting gas content in coal seam. Express Inf Min Ind 7:28–30
4. Wu G, Huang M, Li G (2008) Gas content prediction based on BP neural network. Coal Geol Explor 1:30–32
5. Chang TC, Lin SJ (1999) Grey relation analysis of carbon dioxide emissions from industrial production and energy uses in Taiwan. J Environ Manage 56:247–257
6. Hao B, Li C, Wang C (2010) Application of grey correlation degree in the identification of sources of mine water bursting. China Coal, China, pp 20–22
7. Xu B, Liu Z (2000) MATLAB engineering mathematics and applications. Tsinghua University Press, Beijing
8. Zhao Y (2009) Predicting model of gas content based on improved BP neural network. Ind Mine Autom 4:11–13

Optimization Analysis of Dynamic Sample Number and Hidden Layer Node Number Based on BP Neural Network

Chunyun Xu and Chuanfang Xu

Abstract Taking use of BP neural network theory, a cyclic search method between dynamic training samples number and dynamic hidden nodes number is proposed to improve the prediction accuracy of network model. By increasing the training samples one by one to control the network training, searching dynamic sample error expectation matrix, and finally getting the best training sample number, with the goal of minimum error expectations. Thus, through the optimization model of hidden nodes number, searching the number of hidden layer nodes of network minimum output error. The results of the analysis of examples shows that, with the increasing number of training samples, the network output error expectation experienced three stages, namely recurrent big error stage, yield decline error stage and stable small error stage. But with the increasing number of hidden layer nodes, the result is on the contrary. This shows that proper number of training samples and hidden layer nodes is of great significance to improving the output precision of neural network.

Keywords BP neural network · Dynamic sample number · Dynamic hidden layer node number · Cyclic search · Prediction accuracy

C. Xu (✉)
School of Mechanical Engineering, Anhui University of Science and Technology, Huainan 232001, China
e-mail: cyxcumt@126.com

C. Xu
School of Chemical Engineering, Hefei University of Technology, Hefei 230009, China
e-mail: xubiyun_2008@126.com

1 Introduction

Neural network has obvious advantages for nonlinear model predictive compared with traditional forecasting models. But the structure and parameters of neural network have a great influence on the model prediction accuracy, therefore, how to optimize the network prediction accuracy is the key to use neural network. There has been a lot of literatures [1, 2] studying the influence of network output accuracy, but there is very little research on the influence of network output precision from the sample number and the number of hidden nodes. Cheng Guo's the analysis and optimization of factors affecting the prediction accuracy of neural network model [3] shows that there are three important parameters affecting BP network predictive ability, the input layer nodes, the hidden nodes number and the target value of network training error. Song Shaoyun and others conducted a study analyzing the influence of the artificial neural network sample on output accuracy and reached similar conclusions [4]. Wu Wei and others gave about several comparative studies about the number of hidden units [5]. These literatures mostly describe the impact of network output accuracy qualitatively, not giving concrete quantitative method of the factors affecting network prediction accuracy.

In view of this, this article studies the minimum prediction error of the network training sample number through the method of cyclic search based on BP network which is most commonly used. It helps to reduce blindness of sampling number, especially benefit to some relatively difficult situation of experimental sampling. In other words, it can reduce the cost of sampling under the guarantee of prediction confidence. Previously, the number of hidden nodes is mostly determined by experience formula, getting through the method of trying-to-take, which has certain limitation. This article firstly gives the optimal number of training samples. Then, according to the optimization model of hidden nodes number, it finally draws the concrete optimal number of network hidden nodes.

2 BP Neural Network Prediction Optimization Model

2.1 Input Layer Sample Number Optimization Model

Firstly gets samples in the ways of experiment acquisition or theoretical calculation, and then numbers the samples.

Define $X_{[n]}$ as dynamic number of training samples, and control the network training by increasing the input layer of training samples, at the same time, define $EXP_{[n]}$ and $SSE_{[n]}$ as one-dimensional dynamic training sample error expectation matrix and the sum of squares respectively to record the training error corresponding to the dynamic training samples. At last determine the optimal input layer sample number through the analysis of dynamic error expectations and dynamic sum of squares due to error.

Dynamic number of training samples meets: $X_{[n]} = X_{[n-1]} + 1 \, (n \geq 2)$

Define $X_{[1]} = 1$, and experiments show that the BP network training cannot be conducted when there is only one sample data. Therefore, to meet the training network, $X_{[2]}$ as the initial dynamics of the number of training samples contains two sample data.

When numbers of samples is $X_{[n]}$ error expectation meets:

$$F\left(X_{[n]}\right) = \sum_{i=1}^{n} |T_i - PRO_i| \Big/ (nT_i)$$

Here, T_i and PRO_i are expressed as the theoretical value (experimental value) and the predicted value of the ith sample when the number of training samples is $X_{[n]}$. Thus, dynamic sample error expectation matrix can be expressed as:

$$EXP_{[n]} = \left[F\left(X_{[2]}\right), \ F\left(X_{[3]}\right), \ldots, \ F\left(X_{[n]}\right)\right] \ (n \geq 2)$$

When number of samples is $X_{[n]}$, its sum of squares due to error meets:

$$G\left(X_{[n]}\right) = \sum_{i=1}^{n} W_i (T_i - PRO_i)^2$$

Here, W_i is expressed as the ith sample network error weights when the number of training samples is $X_{[n]}$, default $W_i = 1$. Thus, the sum of squares of dynamic sample can be expressed as: $SSE_{[n]} = \left[G\left(X_{[2]}\right), \ G\left(X_{[3]}\right), \ \ldots, \ G\left(X_{[n]}\right)\right] \ (n \geq 2)$

Optimal input layer sample number meets: $P = min\left(F\left(X_{[n]}\right)\right)$, $P' = min\left(G\left(X_{[n]}\right)\right)$.

2.2 Hidden Layer Node Number Optimization Model

The existing literature [6] gets the number of hidden nodes according to the experience formula: $P = \sqrt{m + n} + \lambda$, m and n are the input and output nodes respectively, λ takes 1–10 constant. This try-to-take method not only increases the workload, it also has its limitations. Literature [7, 8] takes use of the empirical formula:

$Q = log_2 v$, v is defined as the input layer sample number. However, whether these conclusions are applicable to the specific object of study has yet to be further verified. In view of this, this article takes P the optimal number of training samples as the premise to further optimize the number of hidden nodes.

Define $Y_{[n]}$ as dynamic hidden layer node number, and control the network training by increasing the hidden layer nodes number, at the same time, define $exp_{[n]}$ and $sse_{[n]}$ as one-dimensional dynamic node number error expectation matrix and the sum of squares respectively to record the training error corresponding to

the dynamic of the number of hidden nodes. At last determine the optimal number of hidden nodes through the dynamic error expectations and dynamic sum of squares due to error.

Dynamic hidden layer nodes number meets: $Y_{[n]} = Y_{[n-1]} + 1 \ \ (n \geq 1)$

Define $Y_{[0]} = 0$, and $Y_{[1]}$ as the initial dynamics hidden layer nodes number contains one node.

When number of samples is $Y_{[n]}$, error expectation meets:

$$f\left(Y_{[n]}\right) = \left.\sum_{i=1}^{n} |t_i - pro_i| \right/ (n \cdot t_i)$$

Here, t_i and pro_i are expressed as the theoretical value (experimental value) and the predicted value of the ith sample when the hidden layer node number is $Y_{[n]}$. Thus, dynamic node error expectation matrix can be expressed as:

$$exp_{[n]} = \left[f\left(Y_{[2]}\right), f\left(Y_{[3]}\right), \ldots, f\left(Y_{[n]}\right)\right] \ \ (n \geq 1)$$

When nodes number is $Y_{[n]}$, its sum of squares due to error meets:

$$g\left(Y_{[n]}\right) = \sum_{i=1}^{n} w_i(t_i - pro_i)^2$$

Here, w_i is expressed as network error weights when the number of nodes is $Y_{[i]}$ default v Thus, the sum of squares of dynamic nodes number can be expressed as:

$$sse_{[n]} = \left[g\left(Y_{[1]}\right), g\left(Y_{[2]}\right), \ldots, g\left(Y_{[n]}\right)\right] \ \ (n \geq 1)$$

Optimal hidden layer nodes number meets: $p = min\left(f\left(Y_{[n]}\right)\right)$, $p' = min\left(g\left(Y_{[n]}\right)\right)$.

3 Example Analysis

3.1 Data Pre Processing

Examples of sample data come from source of literature [9]. As the input layer nodes have different sample data dimensions, it has to normalize the sample data before network training in order to improve the speed and sensitivity of network training. Meanwhile, it can effectively avoid the saturated zone of function Sigmoid [10]. This paper adopts the normalization formula as follows:

$$\gamma_{nor} = \frac{2(\gamma - \gamma_{max})}{\gamma_{max} - \gamma_{min}} - 1$$

Here, γ_{nor} represents normalized numerical value, γ represents pretreatment numerical value, γ_{max} and γ_{min} respectively represent the maximum and minimum value of the sample data. The normalized data ranges in the interval $[-1, 1]$.

3.2 The Solution of the Model

According to optimization model of the input layer sample number and the hidden layer nodes number and by means of the Mat lab programming, it gets the results about the effect on prediction accuracy of dynamic sample number and the hidden layer nodes number which is shown in Table 1.

Relationships between dynamic sample number and error expectation and relationships between dynamic hidden layer node number and error expectation are shown in Figs. 1 and 2 respectively.

Here conduct training on neural network according to the optimal number of training samples, namely 12 training samples and 7 hidden nodes numbers, acquired from the model solution. And make use of the trained network to inspect the other five groups of samples. Tables 2 and 3 give the error analysis of prediction model training and inspection results respectively.

The inspection results show: it can be seen from Fig. 3 that when the network training reaches to step 107, the network setting gets to optimal accuracy. From

Table 1 Effect on predictive accuracy of dynamic samples numbers and dynamic hidden layer nodes numbers

Dynamic sample numbers	Error expectation (%)	Dynamic hidden layer nodes numbers	Error expectation (%)
2	35.9755	3	2.4328
3	117.9787	4	2.4727
4	76.5296	5	2.7305
5	95.3259	6	2.8153
6	86.8259	7	2.1126
7	95.2149	8	2.6692
8	97.9469	9	2.5265
9	88.8610	10	110.1653
10	18.1339	11	110.1653
11	2.6768	12	110.1653
12	2.3323	13	110.1653
13	3.3617	14	110.1653
14	8.7377		
15	8.4498		
16	10.9839		
17	8.9507		
18	7.4515		
19	5.7514		

Fig. 1 The relationships between dynamic training sample number and error expectation
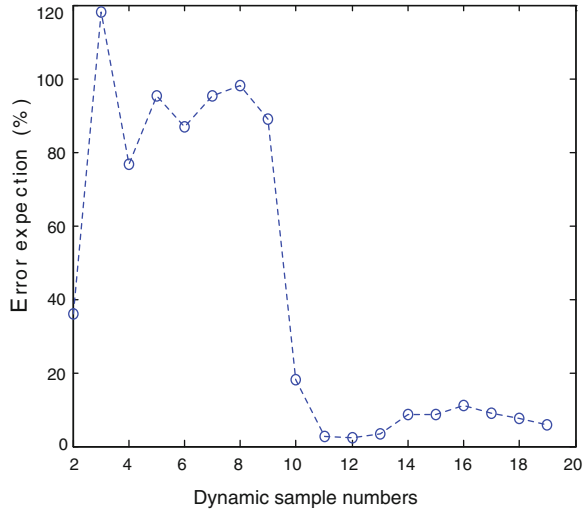
Fig. 2 The relationships between dynamic hidden nodes numbers and error expectation
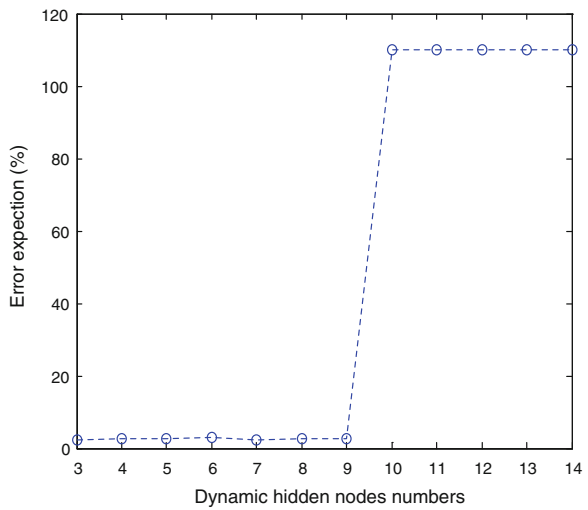
Table 2 Prediction model training results and error analysis

Number of training samples	Average error (%)	Maximum relative error (%)	Relative error is less than 15 % (%)	Sum of squares error (%)
12	8.16	16.98	83.33	5.93

Fig. 4, it can be seen that when $R = 0.99517$, the experimental value has a high degree linear correlation between predictive value, and the model has a good prediction.

Table 3 Prediction model inspection results and error analysis

Number of training samples	Average error (%)	Maximum relative error (%)	Relative error is less than 20 % (%)	Sum of squares error (%)
5	14.86	18.95	80.0	6.00

Fig. 3 The curve of training error
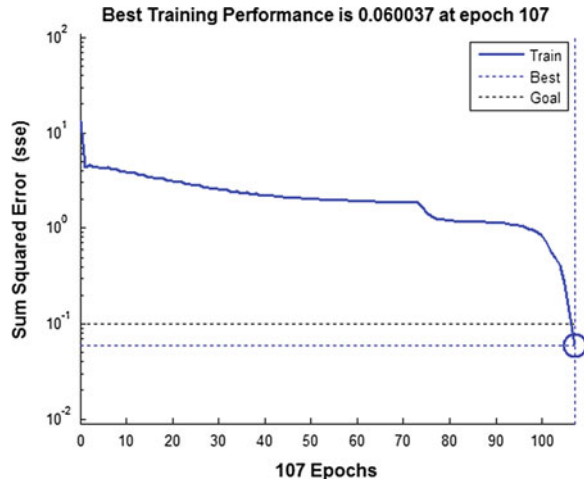
Fig. 4 Regression curve
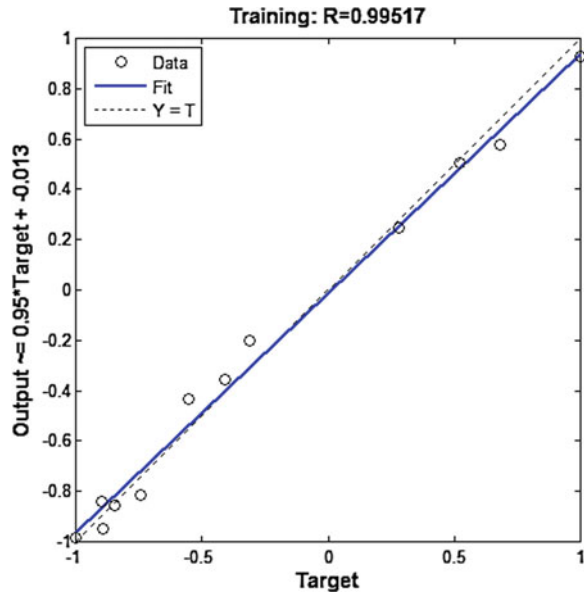
Table 4 Relationship between the range of dynamic sample numbers and the average error

Sample numbers	2–8	9–11	12–19
Average error (%)	86.54	36.56	7.00

Table 5 Relationship between the range of dynamic nodes numbers and the average error

Nodes numbers	3–8	9–10	11–14
Average error (%)	2.54	56.35	110.17

4 Conclusion Analysis

It is known from the network optimization that network output accuracy changes with the change of the number of training samples and the dynamic node number and it changes differently at different stages. As is shown in Table 4, for a dynamic sample set containing18 samples, when the number of samples is in 2–8, the error is relatively large; when the number is in 9–11, the error decreases rapidly; when the number is in 12–19, the error is relatively the smallest. Thus it can be seen that the best range of network training samples is the range from 12 to 19. Table 5 is obtained when the optimal sample number is 12, when the number of nodes is in 3–8, the error is relatively small; when the number of nodes is in 9–10, error rises quickly; when the number of nodes is in 11–14, the error is relatively the greatest. Visibly, the best range of network training nodes numbers is the range from 3 to 8.

Conclusion:

a. with the increase of the number of training samples, the network output error expectation experiences three stages: recurrent big error stage, yield decline error stage and stable small error stage.
b. with the increase of the number of hidden layer nodes, the network output error expectation experiences three stages: stable small error stage, yield rise error stage and recurrent big error stage.
c. in determining the training sample number of network, the number of samples are choose as much as possible, but too much will increase the difficulty of neural network training; the number of hidden nodes is arranged as little as possible, but too little will affect the expected precision of network.

References

1. du Jardin P (2010) The influence of variable selection techniques on model accuracy. J Neurocomput 73:2047–2060
2. Jung J-Y, Phil HL (2006) A neural network model for maximizing prediction accuracy in haplotype tagging SNP selection C 2006 In: International joint conference on neural networks, pp 2303–2307

3. Guo C (2005) Analysis of influence factors for forecasting precision of artificial neural network model and its optimizing. J Patt Recogn Artif Intell 18(5):2303–2307
4. Song S, Bai J The Local Impact Analysis in Artificial Neural Networks. J Value Eng1(7): 144–145
5. Wu W, Chen W (2005) Predict stock price based on BP neural network. J Dalian Univ Technol 41(1):9–15
6. Cai F, Yu J (2011) BP neural network based on grey prediction of gas explosion. J Autom 27(5):42–43
7. Mirchandani G, Cao W (1989) On hidden nodes for neural nets. IEEE Trans Circuits Syst 36(5):661–664
8. Wanas N, Auda G (1998) On the optimal number of hidden nodes in a neural network. In: Canadian conference on electrical and computer engineering, vol. 2, pp 918–921
9. Gao Y (2008) Study of tooth root stress calculation method based on neural network. May 2008, pp 39–48
10. Feng Y, Wu B (2012) Study on forecasting typhoon rainfall based on BP neural network. Res Soil Water Conserv 19(3):289–293

Measurement of Concentration and Size Distribution of Indoor PM10

Hongli Liu, Xiong Zhou and Lei Xiao

Abstract Adopting the microscopic observation imaging and digital image processing technologies, this paper researches a measurement method to the concentration and size distribution of indoor PM10. A new image threshold segmentation method based on the genetic algorithm and Otsu method has been put forward, which can obtain the segmentation threshold by the global optimization, and researches a identification algorithm to the main morphological parameters of a single indoor suspended particulate matter such as size, shape coefficient and the fractal dimension, then calculates the concentration and size distribution of particulate matter using data fusion method. The experimental results show that the method has advantages with intuitive, high precision, fast processing speed, easily data statistics, clearly data analysis and stable measuring results.

Keywords Indoor PM10 (PM10) · Mathematical morphology · Concentration and size distribution · Morphological parameters

1 Introduction

According to the survey, the urban residents spend 90% every day in indoor environment. The suspended particulate matters is one of the most important sources of pollution indoor, which is sucked into the human respiratory tract, and deposit in the lower respiratory tract deposition site under the throat. The smaller

H. Liu (✉) · X. Zhou · L. Xiao
School of Automation, Wuhan University of Technology, Wuhan 430070, Hubei, China
e-mail: james97@163.com

X. Zhou
e-mail: zhouxion_422@126.com

L. Xiao
e-mail: xiaolei218@126.com

the particle size, the deeper the location of respiratory tract, and thus the greater the harm to humans, then the higher morbidity and mortality. The concentration of particles determines the inhaled dose of people, and the particle size carries the important physical and chemical information of particle which is related to its deposition, detention and elimination in the human respiratory.

Generally speaking, it is less possibly that the particles with diameter more than 30 μm enter the lower respiratory tract, most of particles with diameters between 10 and 30 μm depositing in the nasal cavity. The particles with diameters between 5 and 10 μm can enter the trachea and bronchi, and even the blood circulation causing the illnesses related to heart and lung dysfunction. It is more possibility that the particles with diameter less than 2 μm deposit in pulmonary lobe, particularly the respiratory bronchiole and alveolar. Compared with the particles depositing in respiratory tract, the small particles depositing to the pulmonary lobe is more difficult to be cleared away. Therefore, the particles more effect on human health is mainly inhalable particulate matter PM10 (diameters less than 10 μm below). The study of the concentration and size distribution of indoor inhalable suspended particulate matter contributes to understand the impact factors to particles, take positive measures to improve indoor air quality, reduce the harm to the human health [1].

The main methods to detect PM10 concentration are weight method, piezoelectric crystal oscillation method, β-ray absorption, light scattering method, and so on. And the main methods to detect particle size are screening method, settlement method, microscopy method and electron induction method. These detection method have their own best application, but there are some limitations such as long measurement time, complex measurement procedure, measurement accuracy being affected by subjective factor. Moreover, Although adopting the same measuring principle, the results of the same sample often vary greatly with different apparatus and operators.

Adopting the microscopic observation imaging and digital image processing technologies, this paper researches a new measuring method of indoor inhaled particulate matters concentration and size distribution. At first, the paper processes the image of particulate matters based on mathematical morphology, then designs recognition algorithm of the morphological parameters which are particulate matter size, fractal dimension, shape factors and so on, finally calculates the concentration and size distribution of particulate matters with data fusion method.

2 Preparation and Treatment of Particle Samples

2.1 Particulate Matter Collection

Sampled by the low flow air sampler, the suspended particulate matters of indoor air are classified by impact cutting machine. Through the constant weight microporous membrane, the particles are absorbed. The membrane to collect the

particulate matter general adopt cellulose acetate membranes or porous organic thin films that consist of acetic acid and cellulose nitrate. In order to make microscopy imaging easier, the membrane adsorbing particles need to be transparented and fixed by acetone vapor generation device.

Indoor particulate matter with particle size of 10 μm below can aggregate. Particle cohesion is inherent nature of particle itself which interfere to measure particle size distribution accurately [2]. Dispersant and exterior forces are often used to disperse particle cohesion. Dispersant can reduce the surface energy of particles, weaken the gravitational between particles, thus ease or even eliminate the cohesion phenomenon of particles. Different particles samples should choose different dispersant, such as water, water plus glycerol, ethanol, ethanol plus glycerin, and so on. This paper uses glycerol acetate as dispersant to disperse the collected particulate matters.

2.2 Particulate Matters Microscope Imaging

Indoor suspended particulate samples are imaged by ORTHOPLAN polarizing microscope. Adjusting Magnification appropriately, CCD imaging equipment samples the images and transmit image data to computer by computer serial port or USB interface, then through image processing software to process and analyze the digital image, we can achieve the detection purposes of the indoor suspended particulate matter [3].

To ensure accuracy of measurement for observing more smaller particles, the different magnification microscope images are obtained in experiment. Pixel size calibration of image must be precise for the accurate measurement of particle. The results of calibration as a scale, its unit is μm. This detection system chooses standard photolithography micrometer yardstick which is a slides. Through measuring the image of the slide under the microscope directly, we can calibrate the images. Additional the image pixel size varies different magnification microscope, so we must calibrate pixel size of each amplifying image in experiments.

3 Complex Mathematical Morphology Edge Detection

Before measuring and recognizing, the particle images must be preprocessed. Image preprocessing includes the conversion of images, image enhanced, image denoising, threshold segmentation and binarization of image, so the objects and the background of image can be separated. For the binary images through the binarization morphological filtering and edge detection of the mathematical morphology, it is convenient to detect morphological parameters of the image and particle concentration and size distribution [4].

The mathematical morphology edge detection methods with multi-scale and multi-structural elements combined morphological operations with set operations. It can achieve a variety of particle image edge detection with the edge detection operator of different scales structural elements.

The morphological dilation and erosion influence the edge of images, and the difference between the dilated and eroded image and the original image is the edge of image, but these operations are less effective for filtering noise. Obviously the filtering efficiency of morphological open operation and close operation is better, because it used the complementarity of dilate and erode. Only related to the concave, convex location of image edge, the process result do not, so the difference of images can obtain concave, convex features of image, but not reflect all edge features of image. The paper researches an advanced complex edge detecting algorithm, which filters noise with pretreatment of open operation and close operation, smooth images by morphology close operation, and then do dilate operation, finally get better image edge by calculating the difference between dilated image and not dilated image [5]. This algorithm can be described by Eq. 2.

$$(M \cdot B) \oplus B - M \cdot B. \tag{1}$$

where $M = (A \cdot B) \circ B$, A is the original noise images.

4 PM10 Morphological Parameters Detection

PM10 morphological parameters include size, size distribution, shape characterization of particles and so on. The particle morphology is important to study the basic information of particles. In order to facilitate the analysis and calculation, at first, we must mark the particles in binary image, secondly extract the edge contour of particles, began to tracking the edge from the first contour points, calculate particle perimeter. Finally calculate each particle area by scanning Label algorithm.

Supposed A as particle area, p as particle perimeter, the particle equivalent diameter d is defined as

$$d = \frac{4A}{p}. \tag{2}$$

Because the same diameter particles may have different shapes, it is isolated to use one parameter to describe the particles character. Therefore, Here introduce the shape factor and fractal dimension to analysis quantitatively and describe particle character [6, 7]. The formula of the particle shape factor F is:

$$F = \frac{p^2}{4\pi A}. \tag{3}$$

Table 1 Particulate size parameter of inhalable particulate

Sample number	Mean (μm)	Standard deviation (μm)	<5 μm (%)	<2.5 μm (%)
1	1.1542	2.3728	98.52	84.76
2	1.4061	3.3088	95.67	82.31
3	1.2533	2.7835	96.35	84.09
4	1.2039	2.5797	97.48	84.36
5	1.7552	4.6452	92.93	80.51
6	1.3065	3.0567	96.09	83.94

Among many calculation methods of fractal dimension are many, we adopt the calculation method of box dimension [8]. Supposing the side length of box is n (n is the number of pixels included in the side length of box), we divide the edge image into the blocks, each of which has the same number of row and column, then gradually scan all the pixels in a box, Count the number of box N_r pixel whose gray value is zero. According this way, changing the value of side length n, the corresponding number of box N_r can be calculated by the following equation

$$\log N_r = -D \times \log n + k \tag{4}$$

where D is particles fractal dimension, k is a constant [9, 10].

Observed by microscope with different rate multiplying factor, the imaging areas are different, so indoor particulate matter volume conversion factors and calibration coefficients are also different. In 16 × 12.5, 40 × 12.5 and 63 × 12.5 multiplying factor, the calibration coefficients were 14.75, 92.21 and 204.91.

Table 1 is the size parameter of inhalable particulate. With particle size range of 0.03 ∼ 10 μm, larger the mean size of particulate, bigger the standard deviation, less the number of PM5 and PM2.5.

Table 2 is morphological parameters of inhalable particulate matters. The mean value of shape factor of is very close to the parameters of standard square particle, and the shape factors of 80 % particles are more than 0.5 μm. With the increase of particles diameter, their shape factor and fractal dimension all increase. The particle fractal dimension is mostly less than 1.6 μm, and the particles with the

Table 2 Morphological parameters of inhalable particulate

Sample number	shape factor			fractal dimension		
	Mean	Maximum absoluter frequency	Percentage of particles at maximum absoluter frequency (%)	mean	Maximum absoluter frequency	Percentage of particles at maximum absoluter frequency (%)
1	1.1190	1.10–1.20	7.12	1.0442	1.02–1.10	63.25
2	1.1213	1.10–1.20	5.69	1.0575	1.02–1.10	67.57
3	1.1493	1.10–1.20	7.60	1.0532	1.02–1.10	60.38
4	1.1897	1.10–1.20	6.93	1.0657	1.02–1.10	61.32
5	1.1325	1.10–1.20	7.32	1.0720	1.02–1.10	65.98
6	1.1230	1.10–1.20	5.97	1.0564	1.02–1.10	62.15

fractal dimension in 1.02 ~ 1.10 μm are majority. These results show that the overall shape of particles tends to block, and the particles with sharp angle, rectangular and elongated is very little. In general, particle shape is regular.

5 Concentration and Size Distribution Detection of PM10

The concentration of indoor particulate matter include the number concentration, total surface area and weight concentration. The number concentration is defined as the number of particles in unit volume air, and the total surface area refers to the all particle's surface area in a certain range of diameter. In the condition of the same volume, the total surface area of small size particles is large, meanwhile the total surface area of large size particles is small, so the total surface area is not only an important index to measure concentrations of particulate matter, but also a very important parameter to research the toxicological mechanisms of particulate matter. Weight concentration refers to the weight of particles in the unit volume air.

5.1 Number Concentration and Size Distribution

Because the reasons of the resolution ratio, single sample in 16×12.5 multiplying factor, can only detect the smallest size above 0.5 μm particles, the particles with diameter less than 0.5 μm can not be detected, therefore, particle detection data must be integrated with that of particles in 40×12.5 multiplying factor and 63×12.5 multiplying factor, to obtain single-sample particle size distribution of particles by the weighted average method.

N_j is the number of particles that detected in a range of the different particle size, so

$$N_j = \sum_{i=1}^{3} W_i \times n_{i,j}, j = 1, 2, \ldots, 20 \tag{5}$$

where i refers to different multiplying factor, W is the and j is the number of interval of particles diameter. The range of particle diameter is divided into 20 interval with less than 0.5 μm, 0.5 ~1.0 μm, 1.0 ~ 1.5 μm,..., 9.5 ~ 10.0 μm.

5.2 Total Surface Area and Size Distribution of Particles

The greater the surface area, the more absorbed toxic substances, the more hazards to health. Moreover, the more complex form of particles, the surface area more greater.

Supposed indoor PM10 as spherical particles, the diameter is d, the formula for calculating the surface area is:

$$\delta = \frac{s}{V} = \frac{4p}{\pi A}.\tag{6}$$

Because the particle image is two dimension, we can not get the diameter of particles precisely, and use estimate method to calculate.

The surface area of particle s is calculated by the following equation

$$s = \delta \cdot V = \frac{4p}{\pi A} \cdot V = \frac{2p}{3A} D_0^3.\tag{7}$$

5.3 Weight Concentration and Particle Size Distribution

In order to get the weight concentration of indoor PM10, we need to calculate the total particle volume in the range of diameter. Because particles with different diameter have different densities, it is necessary to correct the calculating results. The weight of particles M is shown as

$$M = \omega_1 V_1 + \omega_2 V_2 + \cdots + \omega_n V_n.\tag{8}$$

where ω_i is correct coefficient, which is relevant to particle density.

Because the large particles consist of dust and material from machining, and the small particles are derived from the original combustion particles and aerosols, so the particle density is not a constant in the entire particle diameter range, density of small particles is closed to 1 $g \cdot cm^{-3}$, and the density of larger particles closed to the density of 2.5 $g \cdot cm^{-3}$ [11].

Figures 1, 2 and 3 are number concentration distribution, surface concentration distribution and mass concentration distribution of indoor PM10 under three-ratio measurement data fusion respectively. The diameters of all the particles are lesser than 10 μm, the majority of which distribute in the range of 0.5–4.5 μm, and the proportion of PM2.5 is the largest. The mass concentration distribution curve is very close to the volume concentration distribution curve, and the volume concentrations of smaller particles is very small as a result of small mass concentration. The surface concentration and volume concentration with diameters of less than 5 μm are absolute advantage, so these particles increase harmful to human health.

Adopting this new detecting method, the total weight of indoor inhalable particulate matters is 135 $\mu g \cdot cm^{-3}$, meanwhile, the total weight is 143 $\mu g \cdot cm^{-3}$ by the standard weighing method in accordance with the national relevant regulations. The main reason of error is that the smaller inhalable particulate matters can not be detected due to the lower resolution of microscope.

Fig. 1 Accumulative number concentration and size distribution curve of indoor inhalable particulate matters
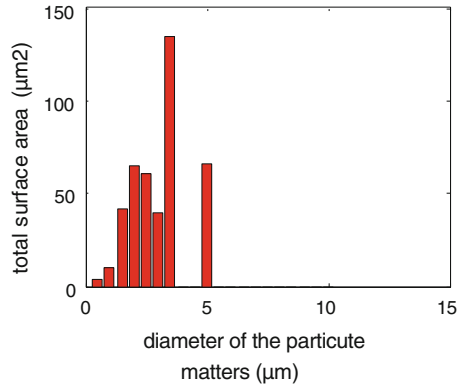
Fig. 2 Total surface area concentration and size distribution curve of indoor inhalable particulate matters
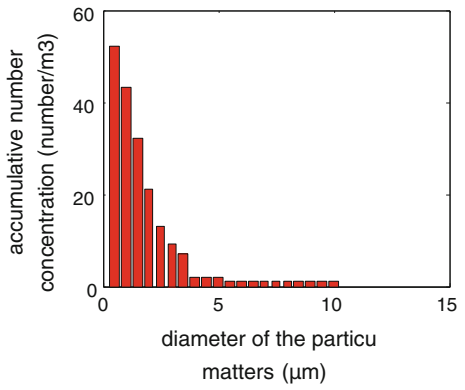
Fig. 3 Mass concentration and size distribution curve of indoor inhalable particulate matters
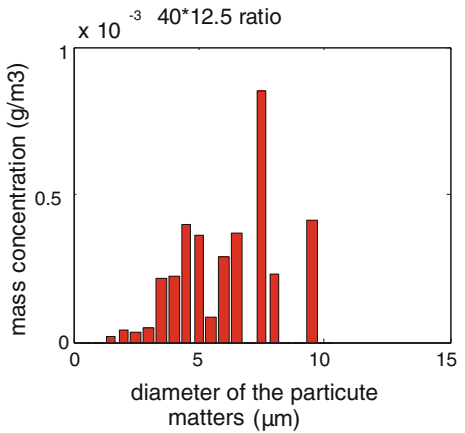
6 Conclusion

Based on microscopic observation and image process, this paper has put forward to a new method to detect the concentration and size distribution of indoorPM10s. This method not only apply the particle morphology and fractal theory to calculate the particle morphology parameters and fractal dimension, also analyses the sources of particles and estimating the density of particles, so can obtain the mass concentration, total surface area and volume concentration and particle size distribution of indoor suspended particulate matters at the same time.

References

1. Naito M, Hayakawa O et al (1998) Effect of particle shape on the particle size distribution measured with commercial equipment. Powder Technol 100(01):52–60
2. Liu H, Li C, Li L, Zhang W (2008) Research of detecting the size distribution for indoor inhaled particulate matters. J Wuhan Univ Technol (Transp Sci Eng) 32(5):884–887
3. Yuan Q (2001) Digital image processing M. Electronic Industry Publishing House, BeiJing, pp 429–480
4. Liu H, Zhao Z, Xie Q, Xu H (2011) Suspending particle edge detection method with multi-scale and multi-structural elements. J Wuhan Univ Technol (Inf Manage Eng) 33(3):346–348
5. Haitao R, Guirong W (2004) The image edge-detection based on mathematics morphology. J Suzhou Univ 20(2):42–45
6. Friedlander SK, Xiong C (2000) Measurments of fractal-likatmospheric particles. Aerosol Sci Technol 31:226–227
7. Cai J, Zhao J, Fang R (2002) Research on the Mensuration of Grain Size Using Computer Vision. J Agric Eng 18(3):161–164
8. Hu X, Zeng W, Wu C (2007) Study on relation between surface fractal dimension and it's particle's shape index. China Powder Sci Technol 13(2):14–17
9. Mosharraf M, Nystrom C (1995) The effect of particle size and shape on the surface specific dissolution rate of micro sized practically insoluble drugs. Int J Pharm 122:35–47
10. Dellino P, Liotino G (2002) The fractal and multifractal dimension of volcanic ash particles contour a test study on the utility and volcanological relevance. J Volcanol Geoth Res 113:1–18
11. Xu K, Sun H, Yang C et al (2006) Application of morphological filtering to online detection of surface crack of medium and heavy steel plate. Chin J Sci Instrum 27(9):1018–1011

A Novel Iris Segmentation Method Based on Multi-Line

Yuhong Jia, Wanjun Hu, Xiaoxi Yu and Miao Qi

Abstract In the iris authentication system, the existence of eyelids and eyelashes noises might generate negative effect on pattern analysis. This paper proposes a novel and accurate iris segmentation method which adequately considers the edges and statistical feature to detect the eyelids and eyelashes noises in the captured iris images. First, multi-scale edge detection is used to get the iris coarse locations, and rank filter is employed to smooth images for determining a more accurate searching area of eyelids. Second, morphological operations as well as line Hough transform are presented to reserve the available edge points for multi-line eyelids fitting. Specially, the adaptive average intensity of individual iris image based on region of interest (ROI) is educed to get the statistical threshold for eyelashes detection. Experimental results indicate that the proposed method can effectively remove the occlusion caused by eyelids and eyelashes, and increase the amount of information (AOI) of segmented iris and improve the iris location accuracy.

Keywords Iris authentication · Eyelid and eyelash detection · Hough transform · Multi-line · Amount of information

Y. Jia
Department of Information Engineering, Jilin Business and Technology College, Changchun 130062, China

W. Hu · X. Yu · M. Qi (✉)
School of Computer Science and Information Technology, Northeast Normal University, Key Laboratory of Intelligent Information Processing of Jilin Universities, Changchun 130117, China
e-mail: qim801@nenu.edu.cn

1 Introduction

Iris authentication as a representative biometrical identification technology has been widely researched. Iris is a ring organ which is between the pupil and the sclera, and contains rich texture information. Besides, iris also has the uniqueness, stability and non-invasive characteristics [1, 2]. Compared with other biometrical identification systems, iris identification has higher accuracy [3–8]. The performance of the iris identification system much depends on the accuracy of the iris location. So we pay abundant attention to segmenting the iris precisely.

The existing iris location methods can be divided into two classes, one considers the ROI which is least occluded by eyelids and eyelashes to represent the segmented iris for authentication, like our previous study [9]. Since amount of information (AOI) of the represent region can not present enough independent characteristics, the higher authentication false is received. The others represents the sophisticated technologies to detect eyelids and eyelashes for segmenting iris finely [10–13], but this is complicated and time-consuming.

Nowadays, the main eyelids and eyelashes detection methods include: Xu et al. and Daugman [3, 4] proposed the circular arc method to detect fluctuation eyelids and statistical method to eliminate eyelashes. He et al. and Tan et al. [14, 15] put up a curvature model and an automation parabolic method to detect eyelids, and proposed a prediction model and the gray threshold value method to deal with eyelashes. Sankowski et al. [16] adapted a circle arc to detect the lower eyelid boundaries, and used a line model to detect the upper eyelid boundaries. Li et al. [17] employed a parabolic integro-differential operation combined with ransac-like techniques to detect eyelids. Pundlik et al. [18] presented a texture measure based on gradients to detect eyelashes. Chen et al. [19] used line Hough transform and gray threshold to deal with eyelids and eyelashes. Most existing eyelids detection methods mainly contain circle template, parabola and Hough transform. Although the high segmentation accuracy has received, the circle template method and Hough transform cost the mess search over the whole image which is a waste of time. And the parabola method emphasizes the accurate fitting points. Moreover, most eyelashes detection methods either greatly reduce the available information of the iris, or only parts of noises are removed, and make the algorithms more complex.

In this paper, a new accurate segmentation method is proposed. First, multi-scale edge detection is considered to get the course eyelids region. Second, line Hough transform and morphological operations are adopted for multi-line eyelids fitting. Finally, an adaptive method is proposed to get the eyelashes intensity for eyelashes detection. Specially, a new evaluation criterion named amount of information (AOI) is proposed. Experimental results show that our method can segment iris more accurately and get the higher amount of information (AOI).

This paper is organized as follows. Section 2 introduces the proposed method. Experimental results are reported in Sect. 3. Finally, conclusions are given in Sect. 4.

2 Proposed Method

The flow chart of proposed method includes four parts: iris coarse localization, lower eyelid detection, upper eyelid detection and eyelash detection (Fig. 1).

2.1 Iris Coarse Location

The iris coarse location as described in literature [9], localizes iris inner and outer boundary with circle Hough transform and only segments two iris subregions (Fig. 2c, d). In this study, we adopt the method in Ref. [9] for detecting the iris inner and outer boundaries.

2.2 Lower Eyelid Detection

The grayscale image filled in reflections is smoothed with an adaptive threshold bilinear interpolation [15] in the $x - axis$ and $y - axis$ direction. Since most eyelashes may disturb the processing of eyelids segmentation. Fuzzy filter operation is used to smooth images for eliminating most eyelashes.

Step 1: Fuzzy Filter Operation

Validated over the training images, the edge directions of eyelids are almost horizontal, while the ones of eyelashes are almost vertical, thin and dark (Fig. 3a).
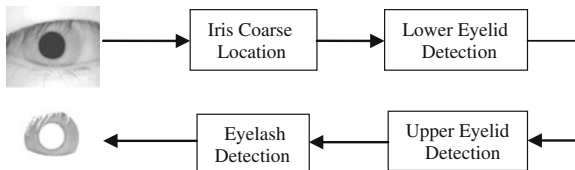
Fig. 1 The flow chart of the proposed method
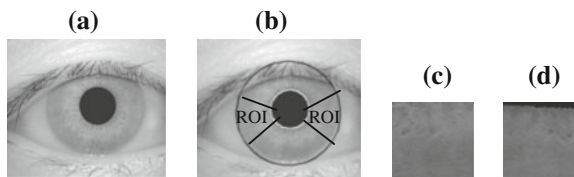
Fig. 2 **a** Original image. **b** Iris boundary labeled image. **c, d** Segmented ROIs
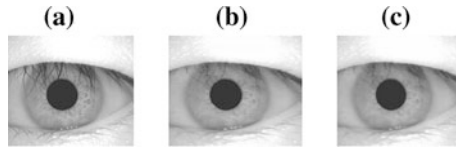
Fig. 3 **a** Original image. **b** Rank filter operation. **c** Median filter operation
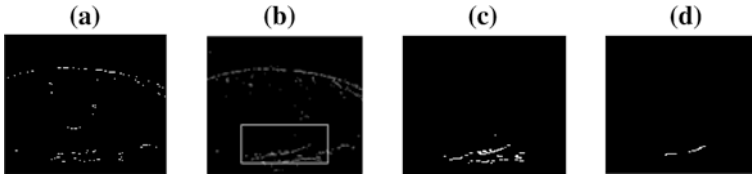
Fig. 4 **a** Edge image via Sobel operation. **b** Edge image with pupil edge removed. **c** Edge image via line dilatation. **d** The gained final lower edge points

A 1-*D* horizontal rank filter (with length $L = 7$ and rank $p = 2$) [14] is adopted to eliminate the most eyelashes, we keep the main contour profile of eyelids (Fig. 3b). Besides, a median filter with sized 3×3 is employed to remove noises caused by the rank filter operation (Fig. 3c).

Step 2: Edge Detection

We use the Sobel edge operation along vertical direction to detect the eyelids edge points (Fig. 4a), and remove the pupil edge points (Fig. 4b).

Then we adopt the line dilatation operation with 1×5 to fill with interspace between the edge points in horizontal direction (Fig. 4c). The longest line is received to gain the valuable edge points for eyelids fitting (Fig. 4d).

2.3 Second Level Heading

Step 1: Edge Detection

Upper eyelids may always be occluded by more eyelashes than lower eyelids. The fuzzy filter operation and edge detection is adopted to get the coarse upper eyelids position. Partial edge points in the rectangle region above the iris centre (Fig. 5b) are reserved. Then we adopt the line dilation operation with 3×20 size to fill with interspace between the edge points in horizontal direction and remove the short lines (Fig. 5c).
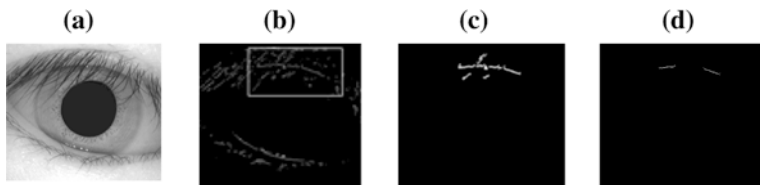
Fig. 5 **a** Original image. **b** Edge image via Sobel operation. **c** Edge image via line dilation. **d** Last upper edge points via Hough transform
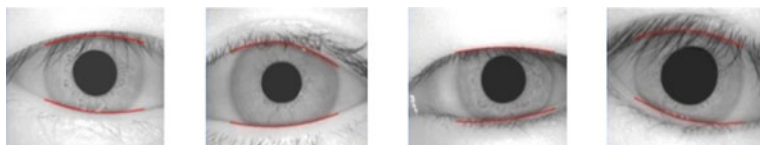
Fig. 6 Partial images labeled with upper and lower eyelids

Step 2: Hough Transform

The upper eyelids could be approximated as parabola curves, and a parabola curve can be expressed as multiple lines (Multi-line) [19]. Hough transform is adopted to keep the valuable edge points.

As the Hough transform is based on voting scheme, it is very tolerant to noise. Line Hough transform is an efficient tool to detect the lines. Therefore, we use the Hough transform to detect much more lines whose Hough peak are not less than a pre-defined threshold. The best result is got when the threshold is 0.7 *Peak*, where the *peak* is the Hough peak value. All the edge points on the lines (Fig. 5d) are remained for eyelids multi-line fitting.

Step 3: Polynomial Fitting

In most cases, the distributions of eyelids edge points are curves. Based on the detected eyelids candidate points, we can fit the eyelids with polynomial fitting:

$$y = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \ldots + \alpha_m x^m \tag{1}$$

$$\theta = \sum_{i=1}^{N} \left(y_i - \sum_{j=0}^{M} \alpha_i x_i^j \right)^2 \tag{2}$$

where $(x_i, y_i)(i = 1 \ldots N)(M < N)$ is the eyelid edge point. Equation (1) is an objective function that enables Eq. (2) to get the minimum variance. The final eyelids curves are shown in Fig. 6.
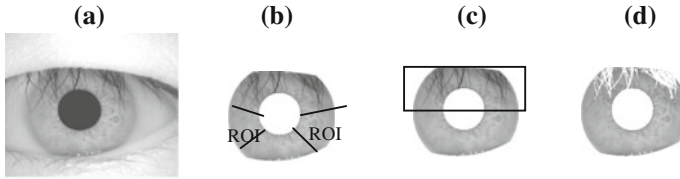
Fig. 7 **a** Original image. **b** Coarse segmented image labeled with *ROI*ₛ. **c** Coarse segmented image labeled with upper region. **d** Final result of iris segmentation

2.4 Eyelash Detection

For accurate iris recognition, it is essential to detect eyelash regions and remove them for iris code generation. In this paper, a new adaptive method is proposed:

Step 1: The pupil, eyelids and the sclera are removed (Fig. 7b).
Step 2: Only two regions of interest (ROI_s) [9] between -35 to $+10°$ and $+175$ to $+210°$ of iris ring are detected (Fig. 7b).
Step 3: The distributions of abundant iris points fit Guassian module, and calculate the average intensity μ and standard deviation θ of (ROI_s):

$$\mu = average(ROI_s) \tag{3}$$

$$\theta = \sqrt{\mu - ROI_s(i,j)^2} \tag{4}$$

Step 4: Finally we only deal with the upper region always occluded by eyelashes, the pixels less than $\mu - \theta - 35$ are considered as the eyelashes (Fig. 7c, d).

3 Experimental Result

To evaluate the performance of our proposed method, we apply it on CASIA Iris Image Database (version 1.0) [20]. This database is composed of 756 iris images 320×280 acquired from 108 different eyes, 7 different images per eye, at monthly intervals. The proposed iris segmentation method can segment the iris effectively, and even the iris is occluded seriously by eyelids and eyelashes. Figure 8 shows partial iris segmentation results. We evaluate the performance of the proposed iris segmentation algorithm from following three aspects.
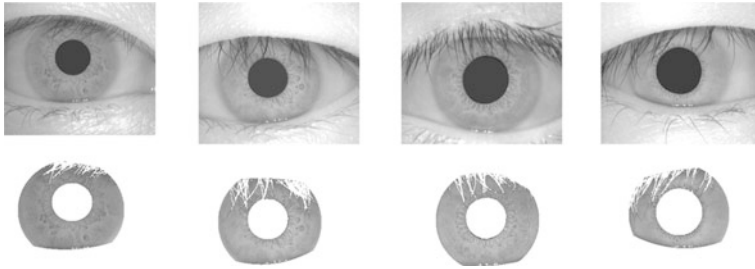
Fig. 8 The partial iris segmentation results of the CASIA iris database

Table 1 Eyelids detection accuracy results on CASIA database

	Upper eyelids (%)	Lower eyelids (%)
Ref. [21]	82.17	93.39
Ref. [9]	91.33	98.45
Ref. [19]	94.76	96.69
Proposed	95.83	98.75

3.1 Eyelid Detection Accuracy

Compared with traditional methods described in Refs. [9, 19, 21], our proposed method gets the higher eyelids detection accuracy listed in Table 1. The proposed approach is increased by 13.66 and 4.50 % for the upper eyelids location, and enhances the accuracy of lower eyelids location by 5.36 and 0.30 % against Refs. [9, 21]. Besides, our method gets higher upper and lower eyelids location precision by 1.07 and 2.06 % than Ref. [19].

3.2 Eyelid Detection Time

The eyelid detection method in Ref. [19] is described as follow, after the first line is detected, the edge points on that line are removed as well as all edge points within 5 pixels above and below that same line. The line Hough transform is then applied again, repeat the process until the line Hough transform peak is less than *Router*/4, where the *Router* is the outer radius of the iris. It needs about average 1.4918 s to locate eyelids. The proposed method just costs average 1.2638 s (Table 2).

Table 2 The processing time of eyelids detection using all the iris images of CASIA DB

	Min(s)	Max(s)	Mean(s)
Ref. [19]	1.2500	2.4688	1.4918
Proposed	1.0469	2.8734	1.2638

3.3 Amount of Information Detection

We propose an innovative evaluation criterion called amount of information (AOI) defined as Eq. (5,) who goes through experimental verification, and represents the proportion of the correct segmented iris corresponding to the groundtruth, and all the groundtruths are obtained manually. The more *AOI* we get, the higher iris segmentation accuracy will be received. Besides we adopt noise detection rate η, iris discrimination rate ζ and average discrimination *Mean* presented in Refs. [22, 23] as evaluation metrics:

$$AOI = \frac{TP_I}{N_I}, \; \varsigma = \frac{TP_I}{TP_I + TN_I}, \; \eta = \frac{TP_n}{TP_n + TN_n}, \; Mean = \frac{\xi + \eta}{2} \qquad (5)$$

where subscripts η denotes noise (including eyelids and eyelashes) and I denotes iris respectively. TP_n and TP_I denote the numbers of noise points and iris points detected correctly. TN_n and TN_I denote the numbers of noise points and iris points detected incorrectly, and N_I denotes the numbers of iris points in the groundtruth.

Table 3 and Fig. 9 list the comparison results of the proposed method with Refs. [9, 19]. We can see that our proposed approach achieves the higher mean and *AOI* detection rate. The proposed approach is increased by 30.34 % for average detection accuracy and 64.60 % for the *AOI* detection results relatives to Ref. [9],

Table 3 Iris segmentation results comparison using all the iris images of CASIA DB

	ς (%)	η (%)	*Mean* (%)	*AOI* (%)
Ref. [9]	91.92	31.98	61.95	32.02
Ref. [19]	86.21	91.73	88.97	93.07
Proposed	91.75	92.83	92.29	96.62

(a)   **(b)**   **(c)**   **(d)**   **(e)**   **(f)**
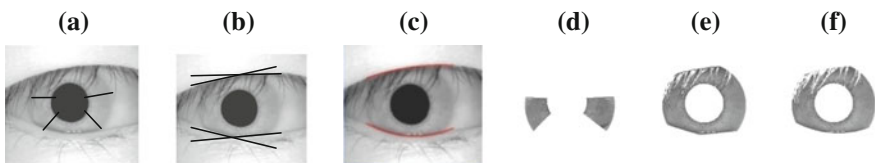
Fig. 9 **a** Iris detection without eyelid detection within ROI_s of Ref. [9]. **b** Eyelid detection result of Ref. [19]. **c** Eyelid detection result of the proposed method. **d** Segmented iris in Ref. [9]. **e** Segmented iris in Ref. [19]. **f** Segmented iris in this paper
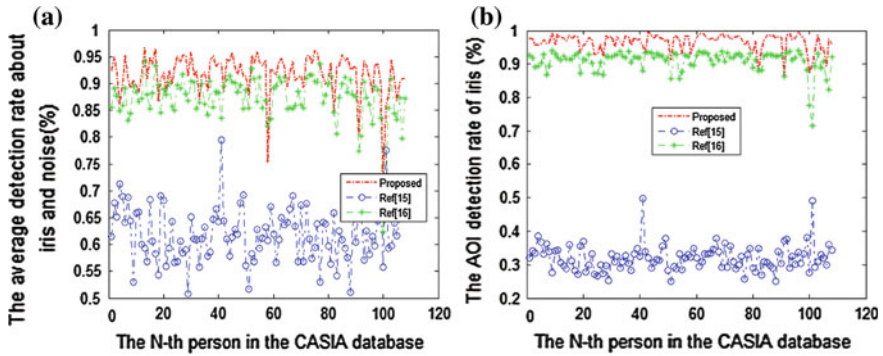
Fig. 10 **a** The iris average detection rate. **b** The AOI of iris detection rate

and enhances the accuracy of average segmentation accuracy by 3.32, and 3.55 % for *AOI* detection results relatives to Ref. [19].

All the iris images in the CASIA database are chosen to analyze the average iris detection rate and the *AOI* detection rate shown as Fig. 10. The *x* coordinate in Fig. 10 represents the *N-th* person in the CASIA database, and the *y* coordinate in Fig. 10a indicates the average detection rate, and the *y* coordinate in the Fig. 10b represents the *AOI* detection rate. Obviously, our approach has higher iris and *AOI* detection rate than the ones in Refs. [9, 19].

4 Conclusion

In this paper, a robust eyelid and eyelash detection approach is proposed to eliminate the main noise occlusion from the iris image. The inner and outer boundaries of iris are located using multi-scale edge detection method in our early study. The eyelids are detected with the line Hough transform and multi-line fitting. Then, the adaptive gray intensity threshold is used to detect the eyelashes. Specially, an innovative evaluation criterion called amount of information (AOI) is proposed to estimate the performance of the proposed methods. The experimental results show that the proposed method can achieve higher accuracy of iris location and more amount of information (AOI) than some existing iris segmentation methods.

References

1. Daugman J (2004) How iris recognition works. IEEE TCSVT, pp 21–30
2. Wildes RP (1997) Iris recognition: an emerging biometric technology. IEEE 85(9):1347–1363
3. Xu G, Zhang Z, Ma Y (2006) Improving the performance of iris recognition system using eyelids and eyelashes detection and iris image enhancement. IEEE TICCI. Beijing, pp 871–876
4. Daugman J (2007) New methods in iris recognition. IEEE TCMCB 37(5):1167–1175
5. Bastys A, Kranauskas J, Kruger V (2011) Iris recognition by fusing different representations of multi-scale Taylor expansion. Comput Vis Image Und 115(6):804–816
6. Du Y, Belcher C, Zhou Z (2010) Scale invariant gabor descriptor-based non-cooperative iris recognition. Signal Process 37:1–10
7. Farouk RM (2011) Iris recognition based on elastic graph matching and Gabor wavelets. Comput Vis Image Und 115(8):1239–1244
8. Roy K, Bhattacharya P, Suen CY (2011) Towards non-ideal iris recognition based on level set method, genetic algorithms and adaptive asymmetrical SVMs. Eng Appl Artif Intel 24:458–475
9. Qi M, Lu Y, Li J, Li X, Kong J (2008) User-specific iris authentication based on feature selection. IEEE CSSE, pp 1040–1043
10. Nguyen K, Fookes C, Sridharan S (2010) Fusing shrinking and expanding active contour models for robust iris segmentation. ISSPA, pp 185–188
11. Tajbakhsh N, Araabi BN, SoltanianZadeh H (2010) Robust iris verification based on local and global variations. Signal Process, pp 1–12
12. Jann F, Usman I, Agha S (2013) Reliable iris localization using Hough transform, histogram-bisection, and eccentricity. Signal Process 93(1):230–241
13. Hansen DW, Ji Q (2010) In the eye of the beholder: a survey of models for eyes and gaze. IEEE TPAMI 32(3):478–500
14. He Z, Tan T, Sun Z, Qiu XC (2008) Robust eyelid, eyelash and shadow localization for iris recognition. IEEE ICIP. San Diego, pp 265–268
15. Tan T, He Z, Sun Z (2008) Efficient and robust segmentation of noisy iris images for non-cooperative iris recognition. Image Vis Comput 28(2):223–230
16. Sankowski W, Grabowsk KI, Zubert M, Napieralska M (2010) Reliable algorithm for iris segmentation in eye image. IVC 28(2):231–237
17. Li P, Liu X, Xiao L, Song Q (2010) Robust and accurate iris segmentation in very noisy iris images. IVC 28(2):246–253
18. Pundlik S, Woodart D, Birchfield S (2010) Iris segmentation in non-ideal images using graph cuts. IVC 28(12):1671–1681
19. Chen Y, Adjouadi M, Han C, Wang J, Barreto A, Rishe N, Andrian J (2010) A highly accurate and computationally efficient approach for unconstrained iris segmentation. IVC 28(2):261–269
20. Information on http://www.cbsr.ia.ac.cn/english/IrisDatabases.asp
21. Cui J et al (2004) A fast and robust iris localization method based on texture segmentation. SPIE Defense Secur Symp 5404:401–408
22. Prati A, Mikic I, Trivedi M, Cucchiara R (2003) Detecting moving shadows: algorithms and evaluation. IEEE TPAMI 25(7):918–923
23. Joshi J, Papanikolopoulos N (2008) Learning of moving cast shadows for dynamic environment. IEEE ICRA, pp 987–992

Maxwell Demon and the Integration Algorithm of Clustering-Partitioning

Jinbiao Wang, Famin Ma and Hongwei Huang

Abstract Maxwell Demon Model was proposed in 1871 to challenge the second law of thermodynamics. While ignoring the energy problem, we can find that Maxwell has started the original clustering model, which implements the clustering of quick molecule and slow molecule naturally. In this paper, we proposed a new model called MCP model, which integrated Maxwell multi-Demon and the ants' movements, and implemented the integration of sample clustering and the whole data set partitioning. We tested our algorithm on UCI examples, the results show it behaves well.

Keywords Maxwell Demon · Multi-Demon model · Ants' movements · Integration of clustering and partitioning · UCI

1 Introduction

Maxwell had designed the famous model "MAXWELL DEMON" in the last chapter of ⟨⟨ Theory of Heat ⟩⟩ in 1871. According to the second law of thermodynamics, a container full of air will end of a state that its entropy is the largest, that is to say the macro temperature is the same everywhere, but the micro molecules present Maxwell distributing. Separating the container into two parts m_0 and m_1,

J. Wang (✉) · F. Ma · H. Huang
Department of Computer Science, Technology Civil Aviation University of China,
TianJin, China
e-mail: jinbiaoww@yahoo.com.cn

F. Ma
e-mail: mafamin_2007@163.com

H. Huang
e-mail: hhw6590057@gmail.com

just showed in Fig. 1, there is an eyelet in each separatrice, where there is a demon who can recognise each molecule. By this recognising, the demons let quick molecules move from m_0 to m_1, and slow molecules move from m_1 to m_0. Then situation changed, the whole container's temperature is not the same everywhere, different parts of the container have their own temperature, the whole entropy in the container reduces. While ignoring the debate of energy consuming, we can easily find that the demon model has implemented the clustering of quick molecule and slow molecule. That is what we inspire from Maxwell Demon model.

Maxwell Demon Model behaves well in the dynamic molecules, this is a 2-clustering problem, while considering a normal clustering problem [1], there are four aspects should attract our attention: (1) the static elements problem; (2) n-clustering problem; (3) two important parameters (the number of clusters and the kernel of the clusters); (4) large scale problems.

Maxwell Demon Model behaves well in the dynamic molecules, this is a 2-clustering problem, while considering a normal clustering problem [1], there are four aspects should attract our attention: (1) the static elements problem; (2) n-clustering problem; (3) two important parameters(the number of clusters and the kernel of the clusters); (4) large scale problems.

For the NO.1 problem, we can integrate the ants' movements [2–6], let the ants move the static elements to the front of the demons; and NO.2 problem needs the Maxwell multi-Demon Model to solve it, just like Fig. 2 showing; problems of NO.3 and NO.4 should be discussed deeply.

2 Two Import Ant Parameters

Each machine learning partitioning algorithm needs two parameters: the number of the clusters and the kernel of the clusters. The number of the clusters means the inherent number the data set has, and the kernel refers to the characteristics attribute value in the centre of one cluster. It's hard to identify the two parameters for there is no priori knowledge, and the data set objects are multi-dimension, random and disorderly.

Fig. 1 Maxwell Demon Model
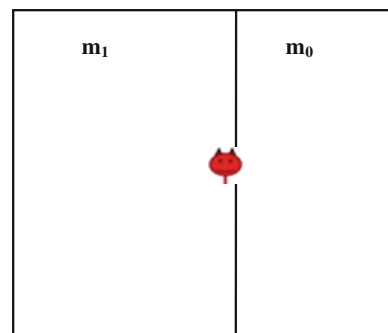
Fig. 2 Maxwell multi-
Demon Model
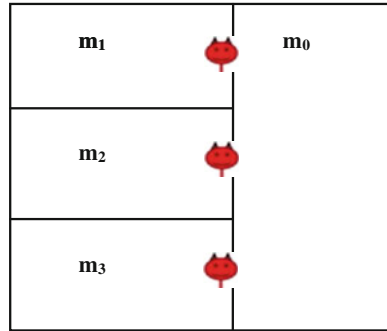
If the number of the clusters is defined by man, the algorithm is called parti-
tioning algorithm, or is defined automatically is called clustering algorithm. The
integrating algorithm of clustering-partitioning refers to the process which at first
finds the number and the kernels of the clusters by sample copy, and then separates
all the objects of the whole data set using the obtained knowledge.

2.1 Characteristic Attributes

The precondition of clustering is the existence of the characteristic attributes,
because the characteristic attributes can really define a cluster, and result in objects
clustering.

So at the first step, we should find out the characteristic attributes to reduce
the computing consuming. And this stage can be obtained by the clustering on
sample copy.

2.2 Constructing the Sample Copy

According to the equal opportunity theory, sample copy composed of elements
collected randomly must be sample copy that extracts elements equally from the
whole data set, the only problem remained is to determine the scale of the sample
copy, these scale should guarantee the existence of the minimal cluster.

The Nyquist theorem [8] tells us: if the sample copy B constructed from sample
A contains at least two elements of the minimal cluster $m_{\min}^A$, it will guarantee the
existence of the minimal cluster. Normally, the empirical value is 5–10.

Prove: extracting sample copy B from data set A, $B \subset A$, from the equal
opportunity theory, we can get that:

$$\frac{|B|}{|A|} \cong \frac{|m_i^B|}{|m_i^A|} \cong \frac{|m_{\min}^B|}{|m_{\min}^A|} = \rho \quad 1 \le i \le q \tag{1}$$

m_i^B represents the cluster i of the sample copy B, $m_i^B \subset m_i^A$, and $m_{\min}^B$ means the smallest cluster of sample copy B, $m_{\min}^B \subset m_{\min}^A$.

And the Nyquist theorem tells us the scale of the sample copy:

$$|m_{\min}^B| \ge \varepsilon \tag{2}$$

$\varepsilon \ge 2$, empirically belongs to 5–10, now we can see:

$$|B| = \rho|A| = \frac{|m_{\min}^B| \cdot |A|}{|m_{\min}^A|} \cong \varepsilon \cdot \frac{|A|}{|m_{\min}^A|} \tag{3}$$

and

$$|m_{\min}^A| \cong \varepsilon \cdot \frac{|A|}{|B|} = \varepsilon \cdot \rho^{-1} \tag{4}$$

ρ is the extracting proportion, it's knowable, $m_{\min}^A$ can be inferred from (4). For example, data set WINE has 178 elements, with its random sample copy contains 30 elements, and the $|m_{\min}^A| \cong 2 \times 178/30 \approx 12$. That is to say, the sample copy scale is at least 12 (Fig. 3).

Based on Sect. 2.1, algorithm can get the two important parameters, the number of the clusters is the number of the demons, and the kernel of one cluster is the criterion that the demon has to let an ant move in or leave out. It solved the NO.3 and NO.4 problems in the Maxwell multi-demon Model.

3 Experiment

We have tested our algorithm on UCI data set [7]: WINE.WINE data are the results of a chemical analysis of wines grown in the same region in Italy but derived from three different cultivars. Experiment details are as follows:
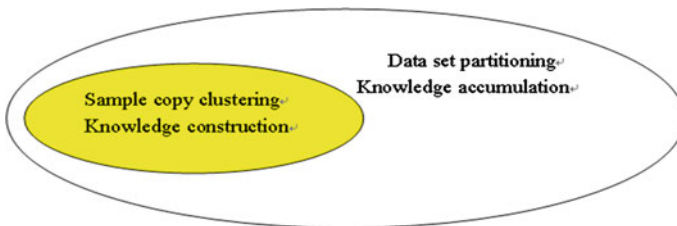
Fig. 3 The integrating of clustering and partitioning

In this example, the number of clusters is 3. The core of the 3 clusters is shown in Fig. 4. Three little demons respectively control the three clustering core which from sub-sample clustering; The clever ant colony respectively estimate the possibility of the elements raised which would pass the 3 little demons, and then take those elements to the little demons who have the most possibility, complete the categories. This process is shown in Figs. 5, 6, 7 and 8.

Each element's cluster number is determined by

$$m_k = \{a_i | \psi_{ik} = \max(\psi_{i1}, \psi_{i2}, \psi_{i3}), 1 \le i \le 178\}, \quad 1 \le k \le 3.$$

Generally, any characteristic attribute significant clustering samples could use this algorithm to process.

If the breaking points are obvious, our algorithm will behave well. The UCI data sets SOYBEAN, LENSES, BALANCE etc. meet the condition. Experiment results are showed in Table 1.

The evaluation system [9] we use includes: Precision, Recall, average evaluation F1 and clustering rate C, their formulas are as follows:

MacroP is the macro precision:

$$MacroP = \frac{1}{q} \sum_{i=1}^{q} \frac{I_i}{|m_i|} \tag{5}$$

P_i is the precision of the cluster i, $|m_i|$ is the number of a cluster elements produced by algorithm, I_i is the right number of elements which are clustering exactly in $|m_i|$. n is the number of clusters.

MacroR is the macro recall:

$$MacroR = \frac{1}{q} \sum_{i=1}^{q} \frac{I_i}{N_i} \tag{6}$$

R_i is the recall value of cluster i, N_i is the inherent number of a cluster elements, I_i is the number of elements clustered into cluster I, n is the number of clusters.

MacroF1 is the macro F1 value:

$$MacroF1 = \frac{MacroP \times MacroR \times 2}{MacroP + MacroR} \tag{7}$$
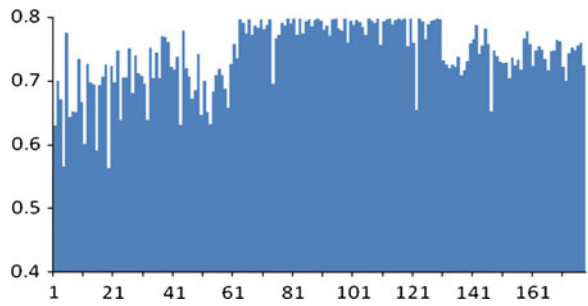
Fig. 4 Average values
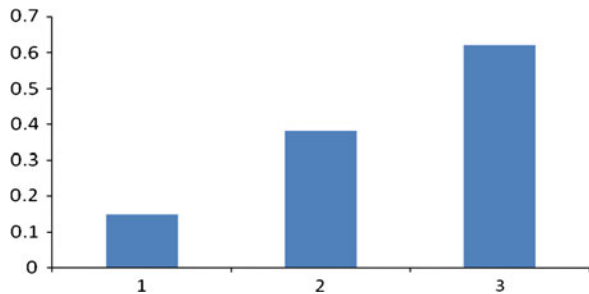
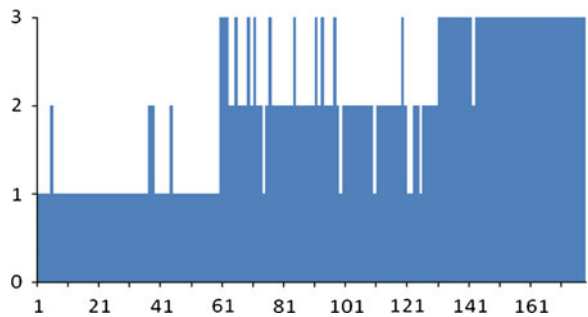**Fig. 5** The distribution of ψ_{i1}
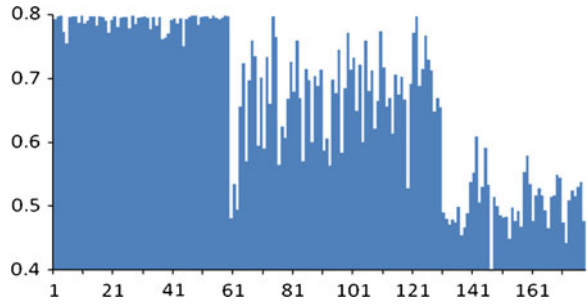
Fig. 6 The distribution of ψ_{i2}
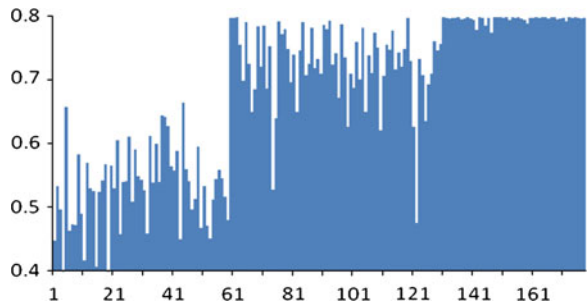
Fig. 7 The distribution of ψ_{i3}

Fig. 8 The distribution of clusters

Table 1 An example of a table

Data set	Algorithm	MacroP (%)	MacroR (%)	MacroF1 (%)	C
Wine	LF	48	88	61	2
178	ATTA	61	84	69	2
13[7,12,13]	MCP	85	87	86	3
Iris	LF	50	89	63	2
150	ATTA	72	82	76	2
4[3, 4]	MCP	88	90	89	3
Ecoli	LF	26	81	39	4
336	ATTA	27	81	40	4
7[2, 7]	MCP	85	94	89	8
Lenses	LF	47	85	59	2
24	ATTA	44	76	55	2
4[3, 4]	MCP	83	93	88	3
Soybean	LF	43	69	53	2
47	ATTA	59	61	59	2
35[23,26,35]	MCP	98	98	98	4
Balancescale	LF	62	58	58	–
625	ATTA	79	12	21	–
4[1–4]	MCP	100	100	100	3

MacroP is the macro precision, and MacroR is the macro recall.
Clustering rate C: represents the number of the clusters that MCP gives out.
When C is more close to 1, the algorithm behaves better.

4 Discussions

Mass data brings in the problem of deeply data mining. How to get the important information from the mass data? On this situation, machine partitioning algorithm was born. The clustering algorithm and the partitioning algorithm is defined by the priori knowledge. Currently, research on clustering algorithm and partitioning algorithm is separated without integrating. Clustering refers to the process of partitioning without any priori knowledge, this is unattached. But as we all know, clustering algorithm is a NP hard problem, this is because at first there are p objects to be chosen, and the next steps are $p - 1$, $p - 2$,…until the clustering ends. So the computing space is:

$$p(p - 1)(p - 2) \cdots (p - q) \approx p! \tag{8}$$

That's the computation complex $O(p!)$.

Although Deneubourg in 1991 had proposed the ant colony clustering algorithm [2], it behaved still not well when dealing with large scale and high dimension problems. Partitioning algorithm which needs multinomial time, it behaves well on large scale problem, the only shortage is that priori knowledge

always involving in human's judging error. In this paper, we constructed Maxwell multi-Demon partitioning model, trying to integrate clustering and partitioning and exert their excellences, as a short, we call this new model MCP. The clustering of sample copy finds its breaking points by sorting characteristic attributes, so the complexity is determined by the sorting computation, an excellence sorting algorithm can complete in $O(p \log p)$. Experiments on UCI data sets show MCP model is feasible. Table 1 also gives the comparison with LF [3], ATTA [4], an improvements comes into sight.

DENEUBOURG's ants participate in our algorithm, but ants in MCP model don't need trivial routes choosing and put all the elements together in spatial space to form a cluster (all they need is a number represents a demon room).Clustering on sample copy characteristic attributes is to get the number of the clusters and the kernels of each cluster, its only operation is to recognize the breaking points, so the performance of new algorithm improved at least 2 magnitudes.

The bottleneck problem in MCP model is the recognition of characteristic attributes. When meeting the data set that has high noise items, the recognition operation needs more time and experts judgements. Another problem is that randomly selection of the sample copy will give rise to different values of the two parameters, the final result will be influenced by its changes, hence, elite strategy should be involved.

References

1. Han J, Kamber M (2001) Data mining concepts and techniques. Machine Press, Beijing
2. Deneubourg JL, Gross S, Franks NR, Sendova-Franks A, Detrain C, Chretien L (1991) The dynamics of collective sorting: Robot-like ants and ant-like robots. Simulation of Adaptative Behavior: From Animals to Animats, 356–363
3. Lumer ED, Faieta B (1994) Diversity and Adaptation in populations of clustering ants. In Proceeding of the 3rd international conference on the simulation of adataptive behavior:from animals to animats 3. Mit Press, 449–508
4. Handl J, Knowles J, Dorigo M (2004) Ant-based clustering and topographic mapping. Artif Life 12(1):35–61
5. Mirzaei G, Jamali MM (2012) Implementation of Ant Clustering Algorithm for IR Imagery in Wind Turbine Applications. MWSCAS, 868–871
6. Zhang C, Zhu M, Zhang B (2012) An improved ant-based clustering algorithm. 2012 8th International conference on natural computation, 749–752
7. Information on http://archive.ics.uci.edu/ml/datasets.html
8. Nyquist H (1928) Certain topics in telegraph transmission theory. Trans AIEE 47:617–644
9. Grira N, Crucianu M, Boujemaa N (2004) Unsupervised and Semi-supervised Clustering:a Brief Survey, A Review of Machine Learning Techniques for Processing Multimedia Content, Report of the MUSCLE European Network of Excellence(FP6)

Image Retrieval Method Based on Hybrid Fractal-Wavelet Image Coding

Haipeng Li, Xu Ji and Jianbo Lu

Abstract This paper presents a hybrid image coding method combined with fractal image coding and wavelet transform. Exploite wavelet transform decomposed an image into some different frequency subbands. We introduced two different processing courses of fractal coding in high frequency subbands and low frequency subbands, obtained corresponding coefficients as index file of an image. Differently with image compression we pay more attention on the accuracy of image retrieval and time consuming. From experimental results we can obtained satisfied results of image matching exploited index file which produced by our proposed hybrid image coding method.

Keywords Image retrieval · Wavelet transform · Fractal image code

1 Introduction

Fractal image coding introduced by Barnsley and Jacquin [1–3] is the outcome of the study of the iterated function system developed in the last decade. The idea of Fractal image coding (FIC) is based on the assumption that the image redundancies can be efficiently utilized by means of block self-affine transformations [4–6]. Because of its high compression ratio and simple decompression method, many researchers have done a lot of research on it. Fractal image compression is widely used in image processing applications such as image signature, texture

H. Li (✉) · X. Ji
Computer Science and Technology Department, Zhuhai College, Jilin University, Zhuhai 519041, China
e-mail: haipengli_jlu@hotmail.com

J. Lu
College of Automation, Northwestern Polytechnical University, Xi'an 710000, China

segmentation, feature extraction, image retrievals. However, this method suffers from a long encoding time as its main drawback. This long encoding time arise from very large number of domain blocks that must be examined to match each range block.

Fractal Image coding can be used to image retrieval based image content. When we extract image feature, actually means to finding a method that can represent image with more brief feature code. That is similarly with image compression. The purpose of image compression is reduce or cancel the information redundancy from image, to obtain more brief feature code to represent original image. Some fractal indexing techniques have been proposed in which the fractal codes are used as the image features [7, 8].

In order to solve the problem of generate effective fractal indexing to represent image features, our work involves exploit a hybrid coder with fractal and wavelet transformation to generate indexing file used in image retrieval. The remaining part of the paper is as follows: Sects. 2 and 3 include Fractal Image Coding and hybrid Fractal Image Coding based Wavelet. Section 4 explains the metric measure between images and experimental results. Section 5 includes conclusion.

2 Fractal Image Coding

Fractal image compression, which is based on the iterated function system (IFS), is an attractive approach applying on image coding. In image compression which exploit fractal encoding, an image is partitioned into non-overlapping blocks, named range blocks, and then approximate each range block with a contractive affine transformation from a domain block, which are selected from the same image but different size. Fractal coding's main point is finding excellent contractive affine transformation [9, 10].

We assume that an image (N × N) be partitioned into several uniform non-overlapped range blocks, each size is r × r. We define these blocks: $R_1,\ldots,R_n$, for each range block R_i, we finding a transformation W_i to mapping a larger domain block $D_{d(i)}$, usually of size 2r × 2r, $d(i) \in \{1,\ldots,n_{D_i}\}$, in the same image. Blocks R_i and $D_{d(i)}$ can be considered pixel intensity. Wi can be described as follows:

$$W_i(R_i) = s_{s(i)}\tau_{\tau(i)}AD_{d(i)} + o_{o(i)}I \tag{1}$$

where $s_{s(i)} \in R$, $s(i) \in \{1,\ldots,n_s\}$, denote scaling factor, $o_{o(i)} \in R$, $o(i) \in \{1,\ldots,n_o\}$, is an offset, A is the operator which shrinks the domain block via pixel averaging to match the range block size, $\tau_{\tau(i)}, \tau_{\tau(i)} \in \{1,\ldots,n_\tau\}$ is a permutation that shuffles the pixel intensities in the domain block, and I is the block with intensity 1 at every pixel.

We call W is a contractive affine mapping, so W is specified by the image partition, each range block, scaling factor, offset, permutation, and the domain block [11].

In fundamental fractal coding algorithm, search all the domain blocks for finding an optimal domain block to approximate a range block, waste long coding time. In the searching course, in order to decrease coding time, we usually yield a suboptimal fractal code in which used technique called collage coding. We should finding out a domain block satisfy minimizes the value:

$$E(R_i, D) = ||R_i - W_i(R_i)|| \tag{2}$$

Gain all fractal parameters means have already complete an image fractal coding, all correlative parameters occupy small storage space, so fractal coding can gain very high image compression ratio.

3 Hybrid Image Code

3.1 Wavelet Transform

Mallat [12] proposed conception of multiresolution analysis of wavelet transform, and pyramid algorithm of signal analysis and reconstruction. The proposed pyramid algorithm used to compute wavelet coefficients, which produced digital QMF team representing DWT. Wavelet transform can be exploited in image processing when from one dimension spread to two dimension. Through horizontal and vertical filtering, DWT decomposed original image to four subbands: low frequency subband C_1^{LL}, high frequency subband C_1^{HL} in horizontal direction, high frequency subband in vertical direction C_1^{LH}, high frequency subband C_1^{HH} in diagonal direction. Furthermore, continue decompose subband C_1^{LL}, can obtain other four subband C_2^{LL}, C_2^{HL}, C_2^{LH}, C_2^{HH}, with more lower resolution. Iterate the process course, an original image be decomposed into $3 * L + 1$ subbands with multiresolution levels and different direction, where L denotes the level of resolution. Figure 1 shows that decomposed diagram with L equates 3.

Hence, wavelet transform decomposed image into mixed multiscale compositions, since carried corresponding sampling step-length in spatial or time domain, doing coarse disposal in lower frequency subband, doing minute disposal in higher frequency subband, finally congregate to any detailed part of objects. This attribute possibly be exploited in image compression with compressing redundant wavelet coefficients with requirements of different resolutions.

3.2 Hybrid Scheme with Fractal and Wavelet

Wavelet transform decomposed an image into subimages with different scales and series of optional spatial signals, which can be exploited reconstructing image. It is an analysis method of combination of spatial domain and time domain. Wavelet transforms as a tool for demonstrating the scale invariance of edges; it is obvious

Fig. 1 Schematic diagram of three layers wavelet decomposition
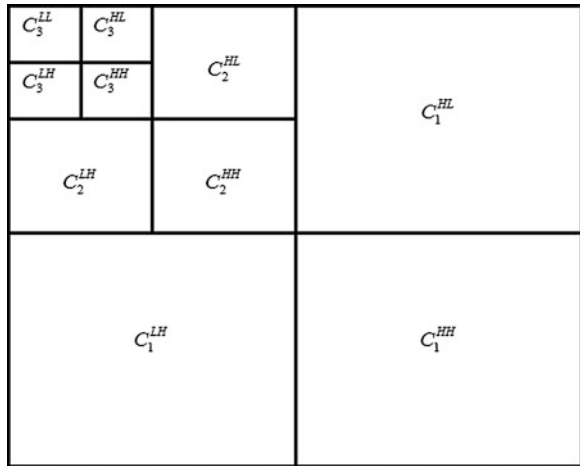
that the subbands in successive wavelet levels are similar. Since the wavelet multiresolution subbands consists of coefficients that have the same spatial location with different scaling resolution, intensity and orientation.

Fractal coding have potential property of high image compression ratio, but neutral images not necessarily have fractal structure so that compression ratio obtained are not steady. Hence many researchers focus on combining different algorithms to overcome the problem of high compression ratio or quality of reconstructing image. We can exploit the scale invariance of edges in wavelet transform to find a form of approximate self—similarity required by IFS.

In [13] proposed the wavelet coefficients have the following relations:

$$Coeff_L^d(R) \approx 2^{-1} Coeff_{L+1}^d(D), \ d = 1, 2, \ and \ 3 \qquad (3)$$

i.e., the coefficients of range blocks at resolution L approximately equals to the coefficients of domain blocks at lower resolution L + 1, multiply by a factor of 2^{-1}.

With this relationship, this paper proposed a hybrid scheme with fractal and wavelet, it is designed to overcome some of the major shortcomings like improving result accuracy and time consuming in image retrieval process.

Our approach of image coding contain two parts: decompose input image into 3 * L + 1 subimages with wavelet transform of L levels. Most energy centralize in low frequency subband C_L^{LL}, which be encoded by fractal coding; To the rest 3L high frequency subbands, we exploited the relationship (3) between coefficients of higher frequency subbands in higher levels pyramid and higher frequency of lower levels to obtain image code.

3.2.1 Hybrid Image Coding Method

1. Split the low frequency subband C_L^{LL} into N non-overlapped range blocks with size of n × n, continue carried wavelet transform on C_L^{LL}, obtained new low frequency subband C_{L+1}^{LL} in next level, we record all blocks with size of n × n in subimage C_{L+1}^{LL} as domain blocks.
2. Let r denote the Gray value of pixels in R, d denote the Gray value of pixels in D, search through all of D to find $D_i \in D$ which minimises the root square error:

$$\varepsilon_i = (\frac{1}{n \times n} \sum_{j=1}^{n \times n} (s_i \cdot d_j + o_i - r_j)^2)^{1/2} \tag{4}$$

3. Find corresponding D to each R, record coefficients $coe_l = (s_i, o_i, (e,f)_i)$, where consists of scaling factor s_i, offset o_i, and $(e,f)_i$ denotes the location of domain block. All above finish process in low frequency subbands.

   Start process in high frequency subbands.

4. Take L level high subbands C_L^{LH}, C_L^{HL}, C_L^{HH} as range blocks RL with size r * r each band of non-overlapping subsquares and L1 level high subbands as corresponding domain blocks D_L with a size 2r * 2r.
5. Filter the Domain Blocks by Gaussian kernel filter with a sums unity in horizontal and vertical direction. For each R_L, search corresponding D_L which minimises the mean root square error:

$$\delta = (\frac{1}{n} \sum_{k=1}^{n} ([s' \cdot D_k + o'] - R_k)^2)^{1/2} \tag{5}$$

6. Repeat steps 4, 5 for the other levels. Record corresponding coefficients as $coe_h = (s', o', t)$, where t denotes the indices for each block in each level.

   Finally, we obtained the input image coding coefficients team $T = (coe_l, coe_h)$, which as an index file of the image in image retrieval processing.

4 Experiment and Results

4.1 Similarity Matching

We assume that there are two same sized image I and image J, define hyb(I) and hyb(J) as hybrid code of image I and image J. According to Sect. 3, we define hyb(I) like this:

$$hyb(I) = T = (\ coe_l,\ coe_h) \tag{6}$$

In order to measure the similarity between two images, we define the hybrid coding distance between image I and image J:

$$
\begin{aligned}
hdis(I,J) = \ & ||hyb(I) - hyb(J)|| \\
= \ & \frac{1}{2} \sum_{i,j=0}^{n} \left(||s_i - s_j|| + ||o_i - o_j|| + ||e_i - e_j|| + ||f_i - f_j|| \right) \\
& + \frac{1}{2} \sum_{k,l=0}^{m} \left(||s'_k - s'_l|| + ||o'_k - o'_l|| + ||t_k - t_l|| \right)
\end{aligned}
\tag{7}
$$

We matching the images through searching similar hybrid coding between two images, we compare hyb(I) with hyb(J) of image J which from image database using (7). After searching all images in database we get minimum value of hdis(I,J), we consider that image I and image J have highest similarity. We exploited the distance between image I and image J as the measure standard when matching images.

4.2 Experimental Results

We apply our approach in image retrieval processing where Daubechies wavelet is employed, assume that there is a query image Q, and image database B. And we have pre-stored hybrid image code hyb(B_i) generated by method mentioned in Sect. 3 as index file of each image in B, i = 1, 2, …, M. M is the number of total images in database B. To checking the performance of image retrieval based on our proposed approach, we do experiments as following steps:

1. calculate hyb(Q) of query image Q with the method mentioned in Sect. 3.
2. calculate hdis(Q,Bi) use (7).
3. sort the results of step 2 from low to high, list top 10 as our final experimental results.

In our experiment, we choice 150 images to build image database. In our image database include different classes image: buildings, people, animals, vehicle, etc. Figuer 2 is the matching result when we specify a query image. Where h denotes the value of hdis(Q,B_i).The image in up-left of Fig. 2 is the query image. We have 25 different input images separately carried out image retrieval processing with our proposed hybrid coding method, and repeat same course with traditional fractal coding method. Table 1 shows the comparing results between traditional image fractal coding and our proposed approach in process of image retrieval. From the results we see that more similar of different images, more smaller of the value of distance between two images, and with the method of hybrid coding can obtained more accurate matching and also have advantage of time consuming.
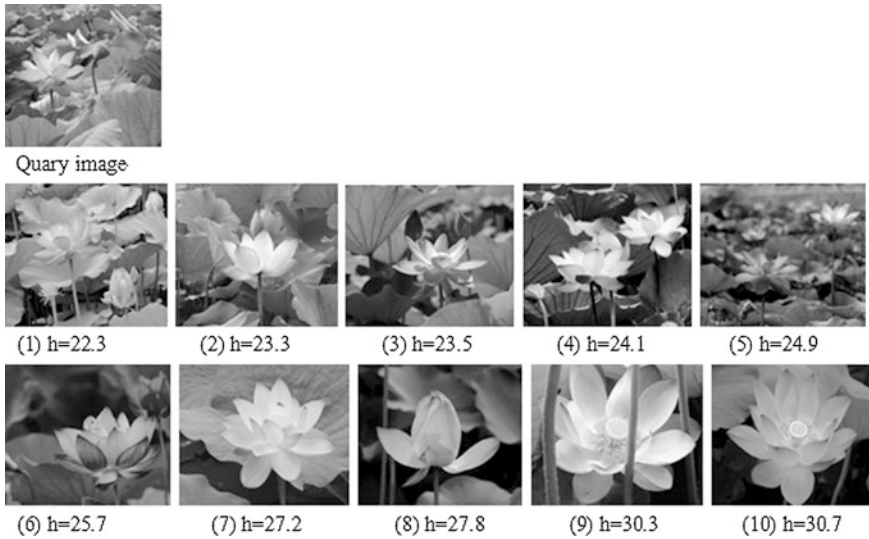
Fig. 2 Image matching results

Table 1 Comparison of image matching results

Scheme	Average matching time(s)	Accuracy (%)
Traditional scheme	64	92
Our proposed scheme	37	98

5 Conclusion

In this paper, we proposed a hybrid image coding method be explored processing of image retrieval. Utilize the relationship of between coefficients of higher frequency subbands in higher levels pyramid and higher frequency of lower levels after wavelet translation, we obtained the image coding scheme in high frequency subimages, and carried out traditional fractal image coding on low frequency subimages. Record the corresponding coefficients as index file of an image, which be exploited in image matching. Through the experiment results, we can see that the proposed method obtained desirable searching results. Detect the performance of image compression utilized the hybrid image coding, and compare the efficiency and performance in image retrieval processing through changing different filters exploited in the wavelet transformation are main points in future work.

References

1. Jacquin AE (1992) Image coding based on a Fractal theory of iterated contractive image transformations. IEEE Trans Image Process 1:18–30
2. Jacquin AE (1993) Fractal Image coding: a review. Proc IEEE 81:1451–1465
3. Barnsley MF (1988) Fractals everywhere. Academic, New York, pp 115–208
4. Barnsley MF, Sloan AD (1988) A better way to compress images. BYTE, pp 215–233
5. Hassaballah M, Makky MM, Mahdy YB (2005) A fast fractal image compression method based entropy. Electron Lett Comput Vis Image Anal 5(1):30–40
6. Chiang JY, Jenq YR (2009) Content-based image retrieval using fractal orthonormal basis. J Inf Sci Eng 25:1255–1266
7. Hsiung SC, Jeng JH (2005) Image retrieval using fractal feature. 9th WSEAS international conference on COMPUTERS (ICCOMP 2005), Athens, Greece, CD-ROM
8. Jeng JH, Tseng CC, Hsieh JG (2009) Study on Huber fractal image compression. IEEE Trans Image Process 18:995–1003
9. Enearncao JL, Peitgen HO (1992) Fractal geometry and computer graphies. Springer, Berlin, Heidelberg, NewYork
10. Bhavani S (2010) A survey on coding algorithms in medical image compression. Int J Comput Sci Eng 02(05):1429–1434
11. Hamzaoui R, Hartenstein H, Saupe D (2000) Local iterative improvement of fractal image codes. Image Vis Comput 18:565–568
12. Mallat S (1989) A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans PAM I 11(7):674–693
13. Khalifa OO, Dlay SS (1998) Wavelets image data compression. IEEE International Symposium on Industrial Electronics, South Africa, pp 407–410

Investment Risk Evaluation of High-Tech Projects Based on Random Forests Model

Guangzhou Chen, Jiaquan Wang and Chuanjun Li

Abstract The investment of high-tech projects is characterized with high risk and huge profit, so a scientific and accurate risk evaluation is of importance to make decision. Based on the analysis of the indicators influencing the investment risk, a typical sample set was chosen according to the quantitative score results from the experts. After training, a random forests regression model was established to comprehensively evaluate the investment risk. The assessment result from a case shows the model not only can give a correct rank to the unknown samples, but also give the contribution degree and importance of the indicators. Subsequently, aimed at the leading model parameters, many tests were done to illustration their impact on the output results of the model. Finally, the importance of variables was analyzed, and was compared with that of projection pursuit model. The above results show it has stronger adaptability and robustness.

G. Chen (✉)
School of Environment and Energy Engineering, Anhui University of Architecture,
Hefei 230022, China
e-mail: chgzh5@163.com

J. Wang
School of Resources and Environmental Engineering, Hefei University of Technology,
Hefei 230009, China

C. Li
School of Management, Anhui University of Architecture, Hefei 230022, China

1 Introduction

At present, it is more and more obvious that the high-tech industry promotes the increment of economy. Now, it has quickly become a new economic growth point. In the investment, because of some uncertain and uncontrolled factors, adding investors' underestimate to the external environment, it becomes more risky to invest in the high-tech projects [1]. Therefore, high risk and high profit are its most remarkable characteristics. In view of the above characteristics, it has been a hotspot to study the risk assessment of investment.

Aimed at the above multi-attribute decision problem, the methods for risk assessment of investment in high-tech projects are various. In the early stage, the scoring evaluation method was frequently employed based on the knowledge of experts. The method was strong subjectivity in deciding the weights of index, so it was not suitable for the large-scale application and evaluation [2]. Wang et al. [3] used the analytic hierarchy process (AHP) to assess the investment risk. However, it was susceptible to subjective factors in constructing the judgment matrix. Jia et al. [4] used fuzzy comprehensive evaluation model and used AHP to identify the weights, so there were also subjective factors in decision. Jia et al. [5] employed multi-level gray evaluation method, and the same shortage was accompanied. With the development of artificial intelligence, artificial neural network (ANN) has been widely applied. In view of its complex non-linear simulation and adaptive capacity, Zhang et al. [2] used it to assess investment risk. Nevertheless, ANN needs a large number of typical training samples for training in order to achieve good generalization [6], or it is prone to result in over-fitting. In addition, how to determine the hierarchy of the network is also a problem.

Recently, some new methods have been proposed. In order to improve the training speed of neural network and evaluation accuracy, Wan et al. [7] proposed a combination evaluation model of Rough Set and BP neural network. Gao [8] proposed a fuzzy AHP (FAHP)-SVM evaluation model; Chen et al. [9] gave a group decision-making evaluation method based on entropy; Chen [10] put forward RS-RBF model; Li et al. [11] employed the projection pursuit model. With the development of statistical theory, the new statistical decision-making technologies have been proposed. Random Forests (RF), proposed by Leo Breiman, is a new classification and prediction model [12]. Random Forests has been widely applied in many fields. The method has the following advantages [13]: (a) fewer parameters to be adjusted; (b) do not worry about over-fitting; (c) suitable to the data set where there are a lot of unknown characteristics; (d) it can be obtained that which features play more important roles in the classification; (e) the data set, where there are a lot of noise, can also achieve better prediction performance. Therefore, in view of the outstanding advantages of the model, the paper imported this method into the risk assessment of high-tech investment.

The rest of this paper is organized as follows. The index system of investment risk assessment for high-tech project is given in Sect. 2. A brief introduction of RF is shown in Sect. 3. Section 4 gives the evaluation result and analysis aimed at the

Table 1 The index system of risk assessment of investment in high-tech projects

R&D, technical risk	Rationality of theoretical basis (x_1), talent resources (x_2), information resources (x_3), condition of R&D (x_4), technical applicability (x_5), technical matching (x_6), life cycle of technology (x_7), technical maturity (x_8), technical advancement (x_9)
Production risk	The level of production equipment (x_{10}), energy, raw material supply (x_{11}), the structure of production personnel (x_{12})
Market risk	Competitiveness force of products (x_{13}), potential competitive effects (x_{14}), marketing capability (x_{15})
Management risk	Rationality of Enterprise organization (x_{16}), science of decision-making (x_{17}), quality and experience of managers (x_{18}), fund supply (x_{19})
Environmental risk	Impact of national industrial policy (x_{20}), macroeconomic impact (x_{21}), natural environment (x_{22}), international environment (x_{23})

impact of model parameters to be set and importance of indicators by a case study. Section 5 summarizes the main conclusions of this paper.

2 The Index System of Risk Assessment of Investment in High-Tech Projects

There are some achievements on the index system of risk assessment of investment in high-tech projects. In practice, some researchers have put forward some risk evaluation index systems, such as that of Mao et al. [14]. After the analysis, he divided the risk of investment into two major categories: systematic risk and nonsystematic risk. Of course, the different indicators were used in different documents, but then, they were basically made up of the R&D, technology, production, market, management, environment, finance risk indicators and other aspects. In order to compare with the existing literatures, this paper uses the factors from the literature [2] to do comprehensive evaluation, and the hierarchical evaluation index system is shown in Table 1 ($x_i, \quad i = 1, 2, \ldots, 23$).

3 The Principle of Random Forests Regression Model

Random forests, proposed by Leo Breiman in 2001, is a new algorithm which is made up of many decision trees. It is formed by the combination of many classification trees, and its classification result is finally decided by the vote law. By the bootstrap re-sampling technique, the method repeats to randomly select n samples with replacement from the original samples set N to generate a new training set, the rest of the cases is used to estimate the error of the tree by predicting their classes. Then k classification trees are generated to form a random forests based on the bootstrapped samples. Finally the classification result of a new

sample is decided according to the score which is obtained by voting from the classification trees.

The method has good generalization performance and accuracy [12]. By combining bagging with randomization method, Breiman realized to reduce the correlation among the classification trees with ensuring the performance of single classification tree at the same time, and the performance of classifiers combined were improved.

4 Case Study

4.1 Building of Training Set and Test Set

In order to conveniently compare the evaluation effects of random forests model, we selected 14 examples of risk assessment of investment in high-tech projects from the economic development innovation center of Fujian province, specific data were shown in the literature [2]. The indicators of risk assessment of investment include both qualitative factors and quantitative factors, finally a total of 23 indicators was selected. After checking and analyzing the business plans submitted by the venture enterprises, experts gave the score of the above indicators to measure the performance of the projects, and then gave final risk scores.

In this paper, ten samples were selected to train the model, the other four samples were tested. They owned high-dimensional and small-sample properties, therefore, the traditional statistical methods were not feasible to use. Aimed at the problem, in this paper we used random forests regression model, which can deal with the samples with high-dimensional, small-sample, redundant and noisy characteristics, to carry out risk assessment of investment in high-tech projects.

4.2 Analyzing of Evaluation Result and Impact of Model Parameters on Result

We selected the above 10 groups of data to train the model, then used random forests regression model in DPS software to perform modeling and calculating importance of variable. In calculation process, we needed to set up three variables: the proportion of training samples, the size of the forest (the number of trees), the number of features randomly selected at the nodes. In the above three parameters, setting up the size of the forest was usually relatively simple, as long as the value was made large enough to ensure that the model could converge. This paper focused on analyzing the choice of the other two parameters.

Taking the scale of the forest as 300 and the number of features randomly selected at the nodes as 5, we carried out many tests under the proportion of

training samples set to 30, 40, 50, 60, 70, 80 and 90 % respectively. The relationship of between the proportion of training samples and mean square error (MSE) of regression for the out of bag samples was shown in Fig. 1. Seen from Fig. 1, taking into account the small differences of each regression calculation, therefore, after up to the proportion of 50 %, we thought that the change from the proportion of training samples was no significant effect on the regression results, and the overall difference was very small.

Under the scale of forest set to 2,000 and the proportion of training samples set to 50 %, we made the number of features which was randomly selected at the nodes increase from 1 to 23 in turn, and analyzed the relationship between the number of features randomly selected at the nodes and mean square error of regression for the out of bag samples. The result was shown in Fig. 2. Seen from Fig. 2, mean square error of regression for the out of bag samples tended to decrease with the increasing of the number of features randomly selected at the nodes on the whole, however, the change is not obvious, and no significant difference.

The experimental results of the two parameters which were set to different values showed that the changes of model parameters had little influence on model regression accuracy, therefore, the model had strong ability to adapt to the change of parameters. In this paper, the model parameters were as follows: the number of random forests was set to 300, the proportion of training samples was set to 90 %, the number of variable features at the nodes was set to 5. After calculating, MSE of regression of the out of bag samples was 0.0791, the final score values of from No. 11 to No. 14 samples were in turn 0.6315, 0.7550, 0.4777, 0.5426, so the rank of the samples was $12 > 11 > 14 > 13$ in decreasing order (that is to say, the score is higher, the risk is lower). In addition, the evaluation results from other methods were shown i n Table 2.

Fig. 1 Performance of model under the different proportion of training samples
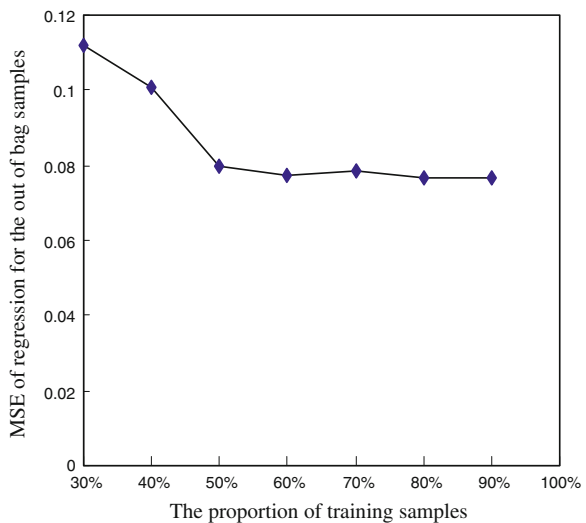
Fig. 2 Performance of model under the different numbers of features
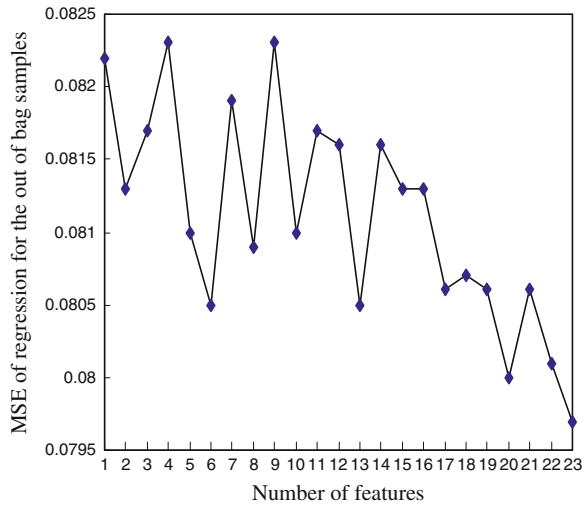
Table 2 The evaluation results of different methods

No. of samples	Rating of experts	BP neural network [2, 6]	Projection pursuit [11]	Random forests
11	2	2	2	2
12	1	1	1	1
13	4	4	4	4
14	3	3	3	3

Seen from Table 2, the evaluation result of random forests regression model was consistent with the results obtained from experts, BP neural networks and projection pursuit model. Therefore, it showed that the model could effectively solve the risk ranking of investment for the alternatives.

4.3 Analysis of Importance for the Variables

The result of each run of random forests regression method was slightly different, thus the importance of some indicators might be a little change, however, the overall trend of variable importance was basically unchanged. The distribution status of importance measure for 23 indicators was shown in Fig. 3, the descending order was as follows: $x_{10}, x_5, x_9, x_{12}, x_{22}, x_2, x_6, x_{13}, x_{20}, x_{18}, x_7, x_{23}, x_4, x_1, x_{19}, x_{14}, x_{17}, x_{16}, x_{15}, x_8, x_3, x_{11}, x_{21}$.

To do a comparison, we took the projection direction of each variable as the vertical axis, and no. of samples as the horizontal axis, then got the distribution status of projection direction for 23 indicators, specific characteristic was shown in

Fig. 3 The distribution status of importance measure for 23 indicators
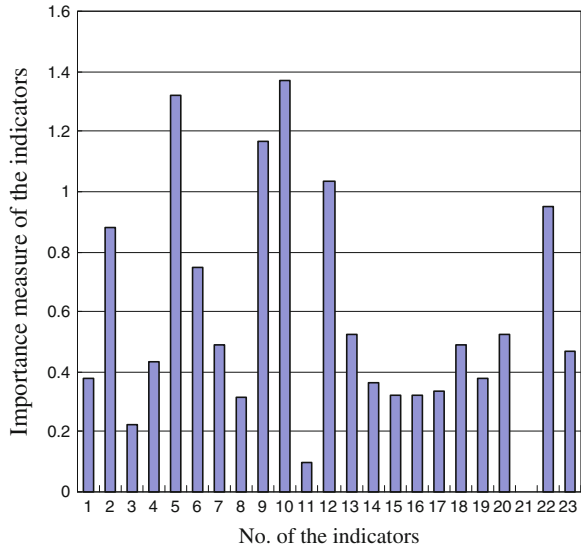
Fig. 4. Figure 4 showed that projection direction of some indicators were negative, so we selected the top 10 important indicators to do a comparison with the result of random forests regression method. The order of the first 10 variables determined by the projection pursuit model was $x_6, x_{22}, x_{12}, x_{18}, x_9, x_5, x_{10}, x_{19}, x_{13}, x_1$. Compared Fig. 3 with Fig. 4, the results showed that although the two methods were a little different for each sort of a specific variable, the importance rank of eight variables in the top ten variables was the same ($x_6, x_{22}, x_{12}, x_{18}, x_9, x_5, x_{10}, x_{13}$).

Fig. 4 The distribution status of projection direction for 23 indicators
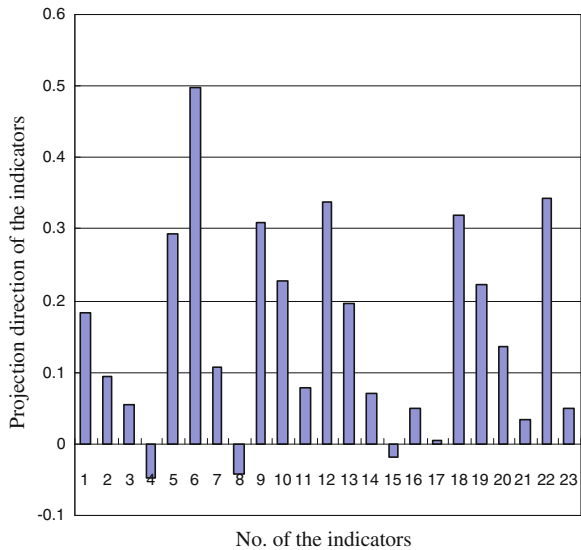
5 Conclusion

For the risk evaluation of investment in high-tech projects, because of many difficulties such as numerous impacting factors, the complex index system and the weights decided difficultly, it is very difficult to make accurate evaluation. For the multi-attribute decision making problem, we built the random forests regression model to assess it. Through the application of the case, the conclusions are as follows: (1) Though evaluation indicators have non-normal, non-linear, noisy, uncertain and other characteristics, the method directly was driven by the training sample data, and fully mined the information of existing data. It can solve the problems that the weights determined difficultly for general evaluation methods, and obtain good evaluation results. (2) According to the importance of indicators, we can make it clear which indicators are the important to assess investment risk so as to provide some reference for decision-making. (3) By the analysis of parameters selected for the model, the result indicates that the model has strong adaptability and robustness.

References

1. Guo MY (2003) Fuzzy comprehensive evaluation of the risks in the development of high-tech products. J Gansu Sci 15(3):118–122
2. Zhang XH (2001) A neural network synthetic evaluation method for the investing risk of the high-technical projects. J China Soc Sci Tech Inf 20(5):608–611
3. Wang SB, Wang SL (2006) Model design of investment risk evaluation for high-tech projects. Enterp Econ 9:26–28
4. Jia XD, Pan DH (2003) Venture capital and investment venture estimation of high-tech enterprises. Inf Control 32(3):204–208
5. Jia XX, Yang ND, Jiang JJ (2004) The multiple gray evaluation on the regional risk of high technology project. Math Pract Theor 34(2):35–41
6. Lou WG (2005) Comprehensive evaluation model for the investing risk of using artificial neural networks in the hi-tech projects. Technol Innov Manage 26(3):8–11
7. Wan HB, Deng YQ, Gao X (2008) Investment risk evaluation model for high-tech projects. Stat Decis 3:63–65
8. Gao YC (2009) Investment risk evaluation for high-tech projects based on FAHP-SVM model. Bus Econ 4:18–20
9. Chen W, Zhang YC, Tian SH (2011) Investment risk assessment of high-tech enterprise based on entropy group decision-making. Sci Technol Prog Policy 28(1):124–127
10. Chen LH (2010) Study on risk evaluation of high-tech projects investment based on RS_RBF. Comput Simul 27(12):207–210
11. Li CJ, Chen GZ (2010) Investment risk evaluation of high-tech projects based on projection pursuit. In: Proceedings of 2010 international conference on the future of the information technology, and management engineering (FITME 2010), Vol 3. Changzhou
12. Breiman L (2001) Random forests. Machine learning 45(1):5–32

13. Tang QY, Feng MG (2007) DPS Data Processing System. Science Press, Beijing
14. Mao JQ, Huo BS, Yang HS (2002) Risk assessment of high-tech project investment. Quant Tech Econ 19(8):37–40

Design and Implementation of Flash Online Recorder

Minhui Wang

Abstract Online recorder software is used extensively because it can acquire data directly and conveniently. More and more online recorder software which base on Flash platform is invented, however, all of these need support of Flash Media Server, it increases expense of development. This paper is trying to achieve local coding and storage by Flash ActionScript3.0 and open source library, supplying a new solution for online recorder.

Keywords Flash · ActionScript3.0 · Online recording

1 Introduction

Windows recorder was always used to recode for a long time. However, there are a lot of limitations on windows recorder especially its recording session. In order to work out the problem about recording session, more and more people decide to install some more professional sound-processing software. Currently, there are some places such as multimedia network teaching platform need sound recording on web page. Sound recording will become more convenient if recorder software is embedded into web page, utilizing convenience of network platform and combining audio acquisition with online network.

Online recorders can open recorded program possibly on web page and upload data to servers automatically. In order to achieve this function, two techniques

M. Wang (✉)
Shanghai Jianqiao University Information Technology College, Shanghai,
People's Republic of China
e-mail: wmh_gench@126.com

should be resolved: (1) Development of clients' recorded program, coding technique contains ActiveX, •Net Window control, Flash (2) Storage and upload of recording files. Application programs of Flash is tiny, doesn't need any install program, because of that, this paper is trying to use Flash platform to achieve program which bases on microphone, ArrayByte and Sound database as3wavsound which all supplied by ActionScript3.0 to record and play recording. Recording files are saved as wav format in local. After embedded into web page, recorders can upload recording files by controls. Some API are not supported under Flash platform because as3wavsound is under Flex platform database, so the paper modifies database in order to transplant the database to Flash platform successfully.

2 Multimedia Audio Technology

Multimedia audio technology is a process which is an exchange between digital data and artificial voice, the process includes: sampling, quantification, quantization bit, frequency of sampling, single and double channel, audio coding and decoding, audio compression format and so on [1]. Sampling is to change successive analog signals to discrete signal on time. Quantification is to change successive analog signals to discrete signal on width. The discrete sample value on width will still be analogous even the value is discrete on time, so the quantity value should be instead by limited PWL level. Quantization bit is a extent of sampling point such as 8, 12 and 16 bit. Frequency of sampling is number of time when analogous waveform of voice changed to digital data, unit of frequency of sampling if Hertz (Hz). According to sampling theory, in order to keep away to distortion, frequency of sampling should be twice time to voice frequency. Single channel is only one sound wave be created when voice is recorded, so does double channel. Higher frequency of quantization bit and sampling, better tone quality it collects.

3 Brief Introduction of Flash ActionScript3.0

ActionScript is executed by AVM of ActionScript in Flash Player. Script's function of ActionScript 3.0 surpasses former ActionScript's. It can establish big databases and complicated programs conveniently. The new-type AVM2 increases its performance. Speed of program executing of ActionScript 3.0 is 10 times than old-type ActionScript's.
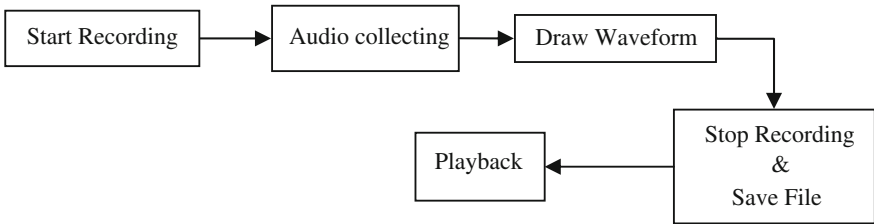
Fig. 1 Process of recording software

4 Design of Software

Before software compiling, plan and flowchart should be drawn at first in order to save time in process, as shown in Fig. 1. The paper achieves online recording and playback, getting rid of dependence of Flash Media Server (FMS), showing wave of microphone when recording.

5 Achievement of Software

5.1 Audio Collecting

The paper is trying to collect voice by MicRecorder class of ActionSctipt3.0's database.

MicRecorder bases on Microphone class in Flash. The Microphone class in Flash is connected with microphone or other kinds of audio input devices which installed on custom system, sending voice data to remote server (as Flash Media Server). Original voice data can be accessed and processed from microphone; Sending voice data directly to loudspeakers or remote server is possible. Coding about voice is completed on servers. In order to get rid of dependence toward Flash Media Server (FMS), as3wavsound database is introduced in database. Partial codes are shown in Fig. 2.

5.2 Coding Storage of Data

The paper will collect wav format audio file which storage bit is a series of 16 bits impulse coding (PCM format).

WAV developed sound file format for Microsoft company. The format follows Resource Interchange File Format (RIEF) filespec, supported by Windows

```
public function MicRecorder(encoder:IEncoder, microphone:Microphone=null,
gain:uint=100, rate:uint=44, silenceLevel:uint=0, timeOut:uint=4000)
        {
            _encoder = encoder;
            _microphone = microphone;
            _gain = gain;
            _rate = rate;
            _silenceLevel = silenceLevel;
            _timeOut = timeOut;
        }
public function record():void
        {
            if ( _microphone == null )
                _microphone = Microphone.getMicrophone();

            _difference = getTimer();

            _microphone.setSilenceLevel(_silenceLevel, _timeOut);
            _microphone.gain = _gain;
            _microphone.rate = _rate;
            _buffer.length = 0;

            _microphone.addEventListener(SampleDataEvent.SAMPLE_DATA,
onSampleData);
            _microphone.addEventListener(StatusEvent.STATUS, onStatus);
        }
```

Fig. 2 Partial codes of MicRecorder

platform and other software extensively. The WAV files has similar format as CD files—44.1 K frequency and 16 bits digit.

Wave files are constructing internal structure according to Resource Interchange File Format (RIFF). The structure of RIFF is dendritic, chunk is its basic unit. Top of the dendritic structure is a "RIFF" chunk, other chunks including "type block identification" (choice), "identifier" (ID), "size of data", "data". Basic forms of WAV file are shown in Table 1.

Write voice data which the double channel is 44.1 KHZ sampling frequency, 16 Bit sampling value to wav format file by WaveEncoder class of as3wavsound.

In main program, addEventListener monitors RecordingEvent. Recording starts when RecordingEvent.RECORDING event is monitored, at this time, self-defining function—onRecording can draw waveform of microphone by Graphics class of Flash, shown as Fig. 3. Graphic class contains a idea to create vector graphics,

Table 1 Basic forms of WAV file

Offset address	Byte count	Data type	Content
00H	4	Char	"RIFF" tag
04H	4	Long int	File length
08H	4	Char	"WAVE" tag
0CH	4	Char	"fmt" tag
10H	4		Transition bytes
14H	2	Int	Form style (10H is voice data of PCM)
16H	2	Int	Channel number, single is 1, double is 2
18H	2	Int	Sample rate (per second), Playback speed of each channel
1CH	4	Long int	Frequency of waveform voice data, the number is channel number × digit per second × digit of each sample

```
public function encode( samples:ByteArray, channels:int=2, bits:int=16,
rate:int=44100 ):ByteArray
        {
            ……
            _bytes.writeUTFBytes( WaveEncoder.RIFF );
        _bytes.writeInt(uint( data.length + 44 ));
        _bytes.writeUTFBytes( WaveEncoder.WAVE );
        _bytes.writeUTFBytes( WaveEncoder.FMT );
        _bytes.writeInt(uint( 16 ));
        _bytes.writeShort( uint( 1 ));
        _bytes.writeShort( channels );
        _bytes.writeInt(rate );
    ……
    }

    private function create( bytes:ByteArray ):ByteArray
    {
        _buffer.endian = Endian.LITTLE_ENDIAN;
        _buffer.length = 0;
        bytes.position = 0;

        while( bytes.bytesAvailable )
            _buffer.writeShort( bytes.readFloat() * (0x7fff * _volume) );
        return _buffer;
    }
```

Fig. 3 Partial code of WaveEncoder class

Fig. 4 Code of onRecording
function

```
function onRecording(event: RecordingEvent):void
{
        line.graphics.clear ();
        line.graphics.lineStyle (1, 0x2AEAEB,100);
        line.graphics.moveTo (5,370);
        activity(recorder.microphone.activityLevel);
}
```

supporting Sprite and Shape object. All these classes contain graphics attribute, it a Graphics object. In order to provide a more convenient usage, assistant functions are supplied: drawRect(), drawRoundRect(), drawCircle() and drawEllipse(). Graphic object can't be created directly by ActionScript code and an exception will be announced if new Graphics() is executed. So, it is necessary to define a Line object of Sprite to draw wave curve [4].

Before using graphic attribute of Sprite object to draw target, lineStyle() method should be identified first, and then starts by moveTo() method to set initial point. Finally, parameter of recorder.microphone.activityLevel is passed by self-defining activity function to achieve dynamic profile.

5.3 Playback of Audio File

In Flash Player 10.1 and AIR 2 edition or even higher, data can be captured as floating point value from microphone. Each value is a sample of single channel. Setting a monitor for sampleData event of Microphone object can get microphone data. Microphone object will use sampleData event to fill buffer of microphone with voice sample. SampleDataEvent contain a data attribute which is a byte array. Unit of sample is floating point values and each value is a sample of single channel.

At recording time, MicRecorder class of as3wavsound database adds sampleData event for Microphone object. Now, SampleDataEvent.SAMPLE_DATA can be monitored and write ByteArray of SampleData by new definition—ByteArray object. Create WavSound and WavSoundChannel object in as3wavsound database and then read ByteArray object. As shown in Fig. 4.

The sound class of Flash itself only permits outside MP3 files to load voice in application. Wavsound is an extensive object of Sound, it can play wav file without Flash IDE or FlexBuilder (Fig. 5).

```
function onRecordComplete(event:Event):void
{
        var player:WavSound = new WavSound(recorder.output);
        var playChannel:WavSoundChannel=new WavSoundChannel(player);
        playChannel.play ();
}
```

Fig. 5 Partial code of recording playback

6 Conclusion

Many companies produce online recorders successively. The paper complies online recording by Flash ActionScript3.0; through introduction and modification to as3wavsound database, audio coding can be achieved without support of Flash Media server; WAV format files can be playback without external player. The new program can be embedded into web page and run by Flash Play1.0 or higher versions, clients can record voice on web page without any kinds of recording software. Currently, this program only supports 44.1 Hz frequency of sampling and doesn't compress these voice data. Because of that, how to compress and transcode voice data successfully under several kinds frequency of sampling by recording software will be the next step of research.

References

1. Windows recorder [EB/OL].2012.9.30.http://zh.wikipedia.org/wiki/Windows錄音機
2. ActionScript3.0 [EB/OL].2012.9.30. http://baike.baidu.com/view/9376.htm#3
3. Hong Z, Pei L, Xine L (2011) Voice acquisition system design and Implementation. Digital Technol Appl 65–66
4. Danpeng S, Yuanhua L (2012) Flash based music visualization description and expression of the application. Sci Technol Inf Appl 143:183

The Intelligent Campus Energy-Saving System Research Based on Power Line Carrier

Jinpeng Li and Guangbin Xu

Abstract This paper introduces a power line carrier technology based on MCU-51 of the campus power saving control system's working principle and structure; At the same time, it introduces direct sequence spread spectrum and half duplex asynchronous modem PL2102 carrier chip, which are used in the power line carrier application system can realize the classroom management and control of the electric light; With PL2102 are given for the control of electric light's basic principle and the method how to realize through software and hardware, along with its design schematic diagram.

Keywords MCU · Power line carrier technology · PL2102

1 Introduction

At present our school campus lighting system is still using simple artificial control mode. This undoubtedly against lighting electricity saving, often there are the light-bright classroom with no one and forgetting to turn off the lights and so on. Incandescent light phenomenon seriously waste; And lamps and lanterns quantity, lighting electricity are difficult to quantify management. So it is necessary for lighting power control mode of innovation, using advanced digital and networked control mode can fundamentally eliminate waste, then it can regulate lighting electricity management [1].

J. Li (✉) · G. Xu
Department of Electrical and Mechanical, Huainan Union University Huainan,
Anhui Province 232001, China
e-mail: 13705542398@163.com

G. Xu
e-mail: hnxuguangbin@126.com

Using advanced electric power carrier communication technology, electric power carrier communication refers to the use of power network as the transmission medium, to realize the data transfer and information exchange of a kind of technology. Power line carrier communication technology is a kind of special communication technology of power system. It provides users with digital community service and electronic business and high-speed Internet access service, voice service, can realize home office and remote home appliance control, realize the fire prevention and guard against theft, prevent poisonous gas leakage and security monitoring system [2]. Power line carrier communication module is the core of power line carrier communication components and hardware foundation, it makes the realization of data, voice, video and power transmission "four networks to one" become possible.

2 System Works

This system is mainly composed of three parts: the host computer, the host, the slave. By pc as the host computer, to the whole system it has the role of monitoring and control. The host is composed of a MCU system and carrier communication module, the slave includes MCU and communication modules and some peripherals, such as some of the A/D sampling unit, address coding unit [3].

The host machine monitors the slave through the way of serial communication of power line carrier. A communication packet includes 8 bytes of data, the first and second bytes are the host identifier; the third byte is command; the fourth and fifth bytes are slave address; the sixth to eighth bytes are data. The host sends command data by radio pattern, the slave machine receive communication package and then analyze data. Firstly, identifying whether the host is their own superior host; Secondly, identifying whether slave address is the own address; When fully confirmed can the slave execute the command and the corresponding operation. The slave machine directly monitor l the state of lamps, through the power line carrier wave receiving circuit receives instructions of the host and executes the corresponding operation; The completion of current, voltage, temperature sampling process, judging lamps are working properly, and according to the data analysis judgment street lamp working condition is normal or not. If the working state is abnormal, implementing the corresponding operation, protecting the lamp, and sending the data and warning to the host.

3 Hardware Circuit Design

3.1 The Main Device Selection and Related Performance

PL2102 is a half-duplex asynchronous modem designed for power line communication network, only by a single +5 V supply, an external interface circuit

coupled to the power line. PL2102 in addition to basic communication control function outside, still built-in 5 kinds of commonly used function circuit: digital frequency correction of the real-time clock circuit, 32 bytes SRAM, voltage monitoring, watchdog timer and reset circuit [4]. They through the standard the I2C interface and external microprocessor are linked together, including real time clock and 32 byte SRAM in the main power supplying power down, can be made of 3 V spare battery power supply to continue work. PL2102 is a low-voltage power line carrier communication chips developed especially for the harsh environment of China's power grid. Due to a direct sequence spread spectrum, digital signal processing, direct digital frequency synthesis technology, and large-scale digital/analog mixed 0.35 μm CMOS process, it has a superior performance in terms of anti-jamming and anti-fading performance as well as domestic and foreign similar products cost-effective.

STC89C52 is 8 bit microcontroller of a low-power, high-performance CMOS with 8 K in-system programmable Flash memory [5]. A single chip has smart 8-bit CPU and in-system programmable Flash to make STC89C52 provide flexible, effective solutions for many embedded control applications. It has the following standard features: 8 k bytes of Flash, 512 bytes of RAM, 32-bit I/O port lines, watchdog timer, built 4 KB the EEPROM, the MAX810 reset circuit, three 16-bit timer/counter, a six-vector two interrupt architecture, a full duplex serial port. In addition to STC89X52 can reduce to 0 Hz static logic operation and support two kinds of software selectable power-saving mode. In idle mode, the CPU stops working and allowing the RAM, timer/counter, serial port, and interrupt system to continue. In power-off protection mode the RAM contents are saved, the oscillator is frozen, stopping all other chips function until the next interrupt or hardware reset. The highest operation frequency is 35 MHZ, 6t/12t optional.

3.2 Circuit Design

The circuit design of the power line carrier is mainly divided into two parts: a power line carrier signal receiving circuit and the transmitting circuit, shown in Fig. 1.

The carrier receiving and sending principle is shown in Fig. 2. Reception signal portion: D7 for clamping, in order to prevent excessive inrush current; C15, C16 and L2 parallel resonance at f = 120 kHz, having the selected frequency of the 120 kHz signal to amplied the input small signal, thereby improving the sensitivity of the reception output. The sending signal portion: the emission loop capacitor C14, the inductor L1 (18 μH/200 mA) for adjusting the emission current and waveform to reduce C14 and increase L1 will reduce the emission current and improve waveform; In the other hand, increasing the emission current and waveform distortion [6]. Due to the coil with load capacity must, adjusting the C14 and L1 will affect coil of the transmitted power and their power. Four transistor step amplification, four diode shield. Emission voltage VHH influence the size of
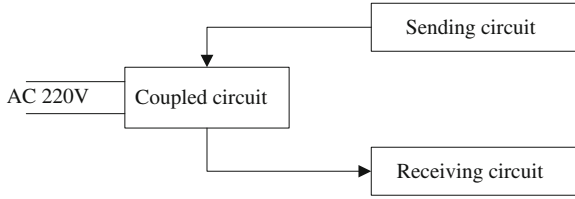
Fig. 1 Carrier transceiver circuit structure
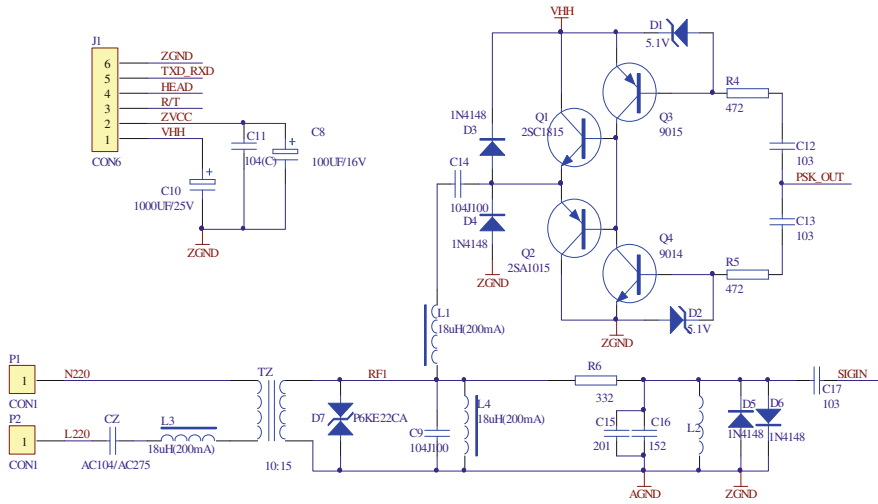
Fig. 2 Carrier circuit schematic

the transmitted power, along with the launch of the voltage drop, emission power decline, launch the higher the voltage of the transistor itself the power consumption, general emission voltage to choose 10–18 V.

The data processing and control circuit are shown in Fig. 3. When the PL2102 received carrier signal processed by the internal circuit, it will demodulate the data bit information, and by RXD_TXD output, while HEAD sync output low. When INT0 of STC89C52 generates an interrupt and receives bit, according to a predetermined format compound the corresponding instruction information. The same time according to a predetermined time in the instruction set to modify the proportion of time of the light switch and lights, and to perform the operation of the light switch, etc. In addition, it will monitor current, temperature, and other parameters after the lights. The implementation of protective measures for seriously overweight and timely failure information is sent to the host [7].
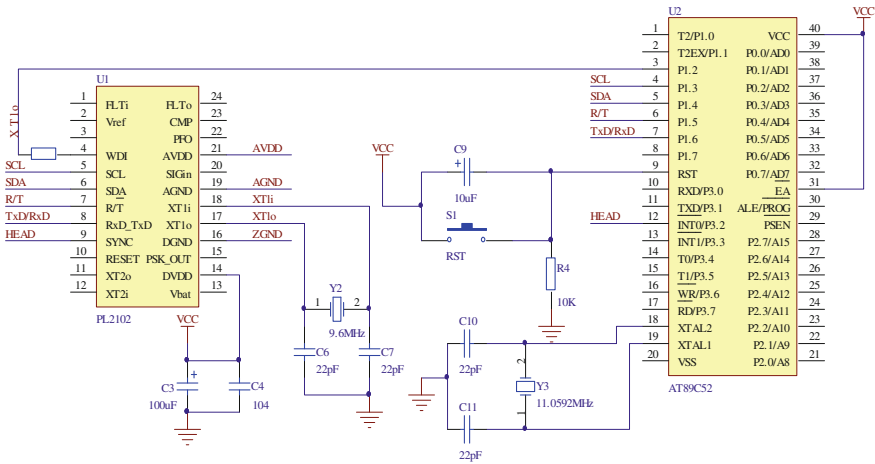
Fig. 3 Data processing and control circuit

4 Software Design

4.1 Carrier Part of the Software Design

System software uses C language, and all modular programming. The part of the software programming how to receive and transmit the carrier signal is the key [8]. Figure 4 show the carrier communication receive process, Fig. 5 is the carrier communication sent process.

4.2 Design of MCU and PC Communication Section

MCU by the way of interrupt communicates with PC, by setting the UART interrupt can control the interrupt control register bits to make the break. MCU sends data process flow chart as shown in Fig. 6.

PC procedure develops in Visual Basic 6.0(VB) environment. VB supports object-oriented programming, own event-driven programming model with structured, and can use the infinite expansion of the control. Using VB, you can use three ways to complete the serial communication [9]. One is provided by using VB communication control with powerful features; Another way is to call WINDOWS API functions, and use the WINDOWS provide strong communications function to write portable applications; The third is done by using file input or output, this method is simple and easy, but there are some limitations.
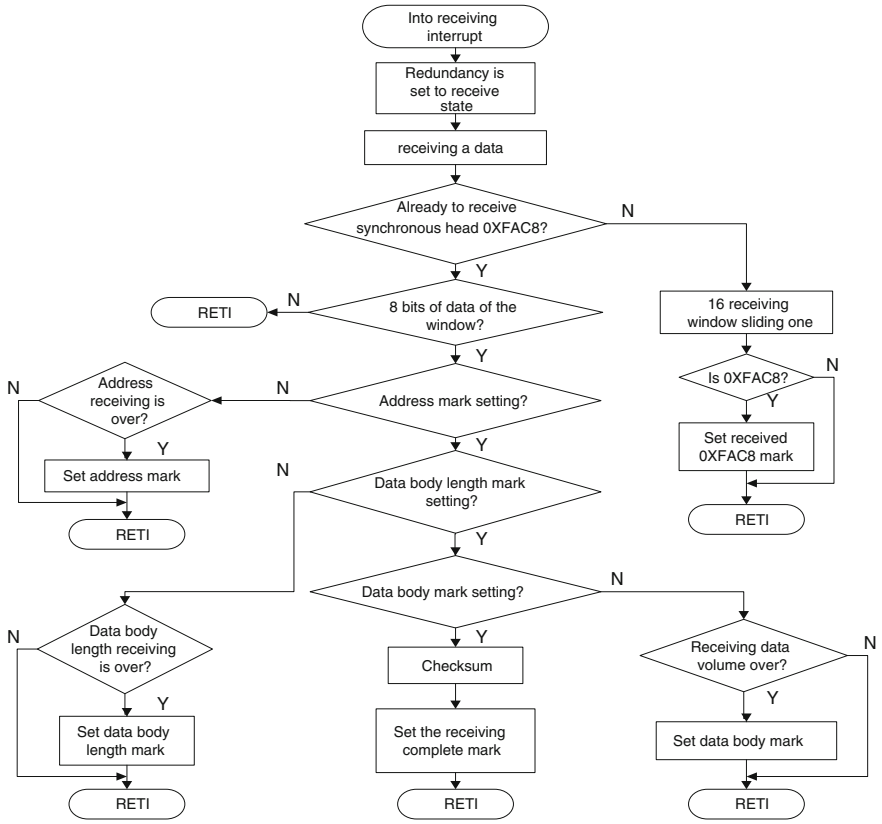
Fig. 4 Carrier communication receiving process

Fig. 5 Carrier communication sent processes
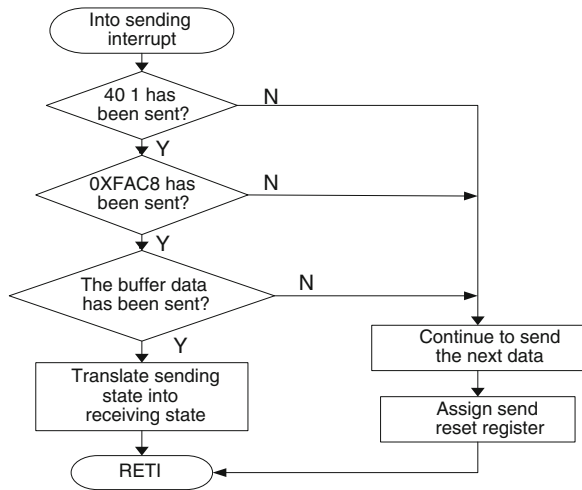
Fig. 6 MCU data transmission program flow
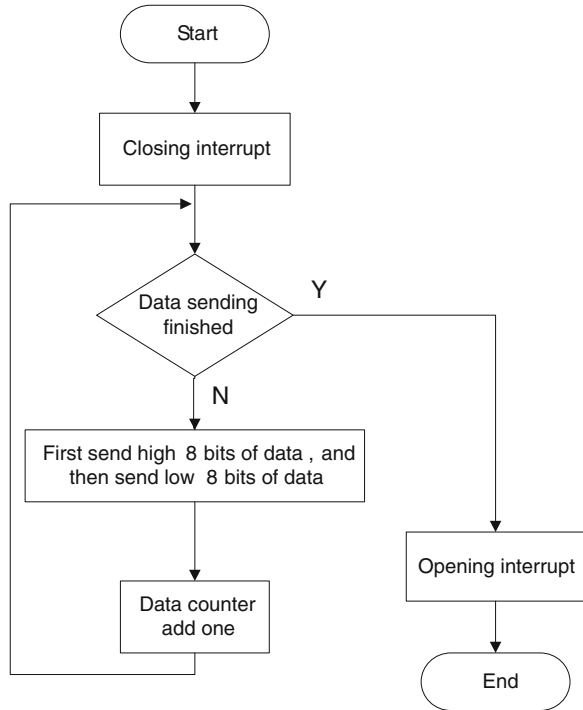
This system adopts the first approach to develop upper machine programs, that is, using VB to provide communication MSComm control, the file named MSComm [10] VBX. This control has complete serial data sending and receiving function, may to the serial interface state and serial communication information format and protocol for setting. Through this control, PC machine available serial port and other equipment to realize easy connection, simply and efficiently realize the communication between the equipments.

5 System Test and Operation

In order to better verify the entire system, on-site installation of the system was carried out in the campus classroom experiments. According to the site's actual operating situation, the system can follow the designed idea for normal work. After nearly a year of running the test, the system works fine, energy consumption reduces by almost 30 %.

6 Concluding Remarks

This major study of power line carrier communication technology and its application of electricity-saving lighting control system on the campus. Through the design of communication module and controller hardware, communication protocols and system software and PC management software, using the existing low-voltage grid transmission signal, and apply it to the lighting control system, system not only improves management efficiency, saves energy and reduces electricity costs, also extends the service life of the lamps, has a very good prospects.

References

1. Ku TT (2009) Identification of customers served by distribution transformer using power line carrier technology. In: 4th IEEE conference on industrial electronics and applications, (ICIEA) pp 3476–3481
2. Twiel MM, Stewart BG, Kemp IJ (2001) Correlation between the bandwidth of PD-detectors and its inherent integration errors. In: Electrical insulation conference and electrical manufacturing and coil winding conference, pp 213–217
3. Foschini GJ, Gans MJ (1998) On limits of wireless communications in a fading environment when using multiple antennas. Wireless Pers Commun 6:311–335
4. Paylidou N, Han Vinck AJ, Yazdani J (2003) Power line communication: state of the art and future trends. IEEE Commun Mag 41(4):34–40
5. Newbury J, Miller W (2001) Multiprotocol routing for automatic remote meter reading using power line carrier systems. IEEE Trans Power Delivery 16(1):1–5
6. Shen CC, Huang Z, Jaikaeo C (2005) Ant-based distributed topology control algorithms for AP HOC mobile networks. Wireless Networks 11(3):299–317
7. Guillemin A, Morel N (2001) An innovative lighting controller integrated in a self-adaptive building control system. Energy Buildings 33:477–487
8. Tsui TSD, Clarkson TG (1994) Spread-spectrum communication techniques. Electron Commun Eng J 6(1):3–12
9. Mathew ST (2012) Portable impedance/reflected power meter for power-line carrier applications. In: Proceedings of IEEE Southeast conference
10. Gu B (2010) The design and realization of acquisition equipment in remote reading system base on low voltage power line carrier. In: Proceedings—2010 international symposium on computational intelligence and design, ISCID vol 1, pp 24–27

Synthetic Decision Support of Broadcasting and Television System

Fulian Yin, Jianping Chai and Jiecong Lin

Abstract To improve the traditional decision support system (DSS) with data driven or model driven only, a synthetic DSS for broadcasting and television system is proposed in this paper. It is consisted by three main bodies which are the combination of data base, model base and method base, the combination of data warehouse and on-line analysis processing (OLAP), and the combination of knowledge base and data mining. The use case of this synthetic DSS is also analyzed in this paper. And the result of typical model proves its validity.

Keywords Broadcasting and television · Decision support system · Data mining

1 Introduction

Decision support system (DSS) was originated in the 1970s by American scientist Keen and Scott Morton. Then it had a great development in the 1980s [1]. Now, it has developed from traditional DSS with data base, model base, method base and knowledge base to a novel DSS which consists of data warehouse [2], on-line

F. Yin (✉) · J. Chai
Information Engineering School, Communication University of China, CUC, Beijing, China
e-mail: yinfulian@cuc.edu.cn

J. Chai
e-mail: jp_chai@cuc.edu.cn

J. Lin
Beijing Institute of Computer Application, Beijing, China

analytical processing and data mining. Its typical feature is to get information which could assist making decisions from a large amount of data [3].

Under the background of "integration of three networks" [4] in china which combines telecommunication network, computer network and cable television network, the broadcasting and television business faces not only huge challenge, but also unlimited opportunities. Business and operation support system (BOSS) in telecommunication has already matured [5] which includes charge and settlement system, operation and account system, client service system and DSS. In the last few years, BOSS for broadcasting and television business is gradually established using the experience of telecommunication BOSS. But, BOSS for broadcasting and television contains only the first three of above systems.

Due to broadcasting and television business, the blank of DSS [6–8] has not been filled. However, DSS becomes more and more imperative with high standard of unified business, scientific decision and efficient management which based on accurate client data. This paper proposes the synthetic DSS of broadcasting and television system which combined traditional data base, model base, method base, knowledge base and novel data warehouse, on-line analytical processing (OLAP) and data mining.

2 The Traditional DSS

2.1 Model driven DSS

Model driven is the main form of traditional DSS through quantitative analysis to assist the decision. The most important original form is based on model base which constructed by mathematical model in management science and operational research. With the development of original DSS and expert system, the mature model driven DSS is proposed. There are three parts with language system, problem processing system and knowledge system as show in Fig. 1.

2.2 Data Driven DSS

Compared with model driven DSS, a new DSS combined with OLAP and data mining is proposed with the development of data warehouse. This is called data driven DSS with the key technique of managing and mining for mass data. Figure 2 gives the architecture of model driven DSS where data warehouse manage the historical data, basic data and integrated data. Through data mining, the spirit of data driven DSS is to obtain the decision information and knowledge through data.

Fig. 1 The architecture of model driven DSS
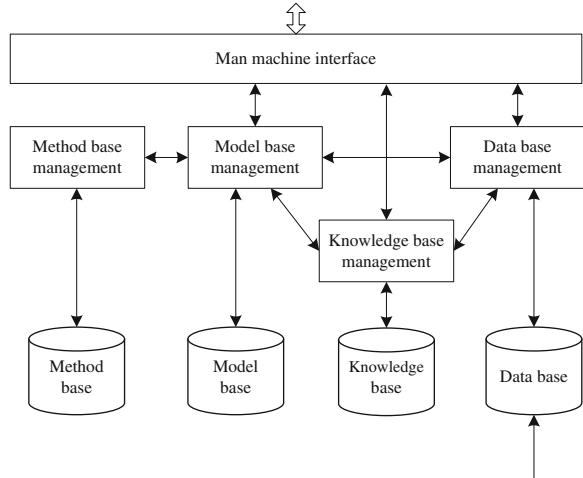
Fig. 2 The architecture of data driven DSS
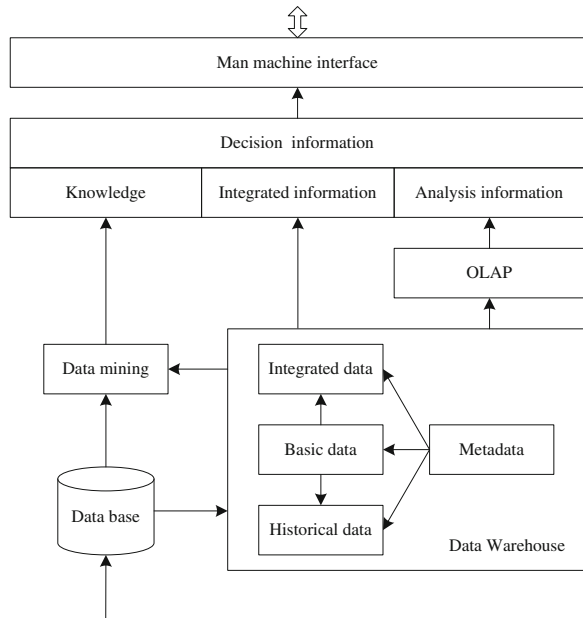
3 The Synthetic Driven Broadcasting and Television DSS

The traditional model driven DSS gives full play to model resource and knowledge resource, besides, realizes the combination of multi models and the storage of a large amount of data. The traditional data driven DSS based on data warehouse and OLAP break a new way for decision support. Wherein, data warehouse and OLAP use data resource to execute decision support and obtain integrated and

predictive information from data warehouse. The model driven DSS and the data driven DSS are completely different. They could not be replaced by each other, but should be combined together.

For broadcasting and television system, a synthetic DSS is proposed in this paper, which combined data base, model base, method base, knowledge base, data warehouse, OLAP and data mining. The new synthetic DSS makes full use of traditional model driven DSS and data driven DSS and could realize more effective decision support. Figure 3 gives the architecture of synthetic DSS for broadcasting and television system. As show in Fig. 3, the man machine interface is the only interaction between decision support system and broadcasting and television clients which include operator clerk, analysis clerk, knowledge clerk and decision maker. Wherein, the decision maker covers all the customers of industry chain like advertisement businessman, programming man, network service provider, and so on.

There are three main bodies for synthetic decision support of broadcasting and television system which are the combination of data base, model base and method base, the combination of data warehouse and OLAP, and the combination of knowledge base and data mining. The above three bodies are both complementary and combinative. Where, the basic component of synthetic DSS is data base which concludes audience background base, audience viewing base and program base. They reflect somebody see some program at some time.
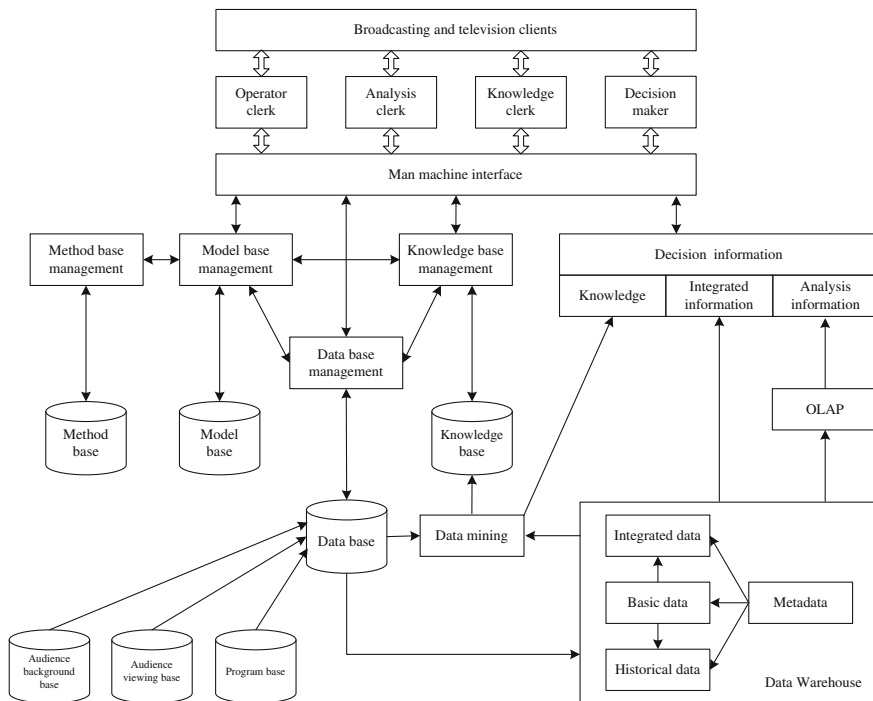
Fig. 3 The architecture of synthetic DSS for broadcasting and television system

The combination of data base, model base and method base is the first part of synthetic DSS. It is an intelligent DSS for broadcasting and television system in order to solve structured problem which has comparatively fixed model. Especially, knowledge is obtain from expert experience or data mining.

The combination of data warehouse and OLAP is the second part of synthetic DSS. For broadcasting and television system, large amount of data is his special characteristic. And the second part obtains integrated information from data warehouse and analysis information from OLAP. Where, model base with mathematic method is added to OLAP in order to enhance the decision ability. Besides, another important decision information is knowledge which obtains from data mining.

The third part is the combination of knowledge and data mining. As show in Fig. 3, it's an important part in synthetic DSS which offers the essence of data to both the first and second part above. Here, the data for data mining from the data base under extracted and the data warehouse composed of metadata.

4 The Use Case of Integrated Broadcasting and Television DSS

The chief function of synthetic DSS is to solve the use case for broadcasting and television clients. To begin with, recognize the problem of clients' use case from man computer interaction. Then, call the corresponding synthetic configuration to analyze the problem. Finally, return decision support suggestion by valuable conclusion, letter report, statistical table or graph. Especially, each use case has its own characteristic and structure. As show in Fig. 4, the use case of synthetic broadcasting and television DSS is divided into structured use case, half structured use case, and non structured use case.

The so-called structured use case is obtained directly from the basic data base and has fixed model to construct. For broadcasting and television system, the structured use case is mainly audience rating index which using the audience viewing base and the program time of program base. Such as audience rating, audience arriving rate, audience operating rate, and so on.

For half structured use case, there exists a certain rule to follow but couldn't completely solving. Here, audience rating preference analysis, audience background analysis, audience rating trajectory analysis and program style analysis are four typical half structured use case which using the result of audience rating index and other basic data base.

Non structured use case has no fixed rule to follow and its determination is hold by the wisdom offered by synthetic DSS. Figure 4 gives some typical ones of non structured use case like characteristic recommendation, audience flow analysis, crowd mining, and so on.
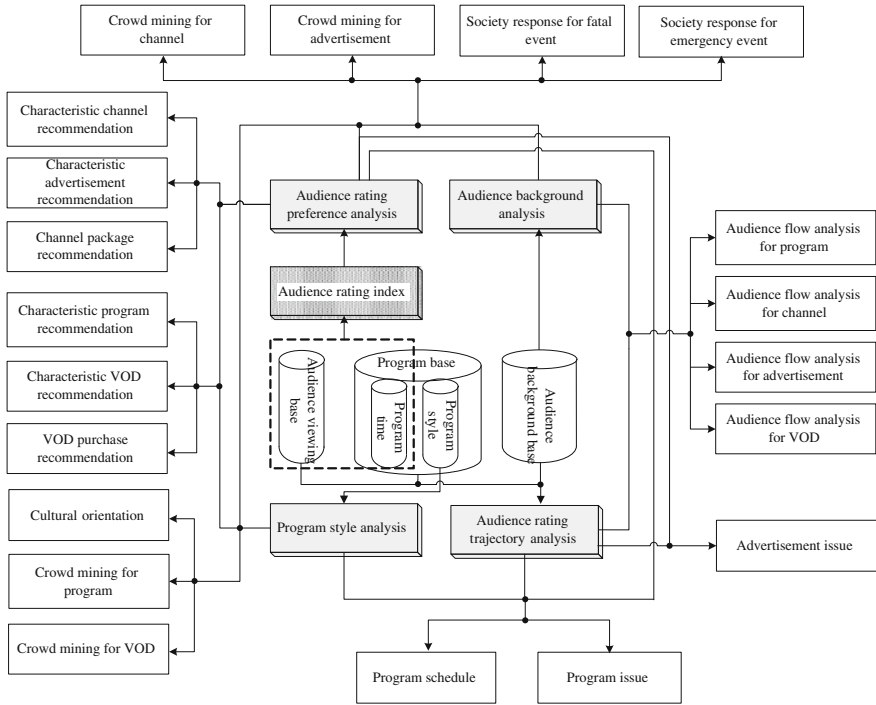
Fig. 4 The structure of typical use case rating
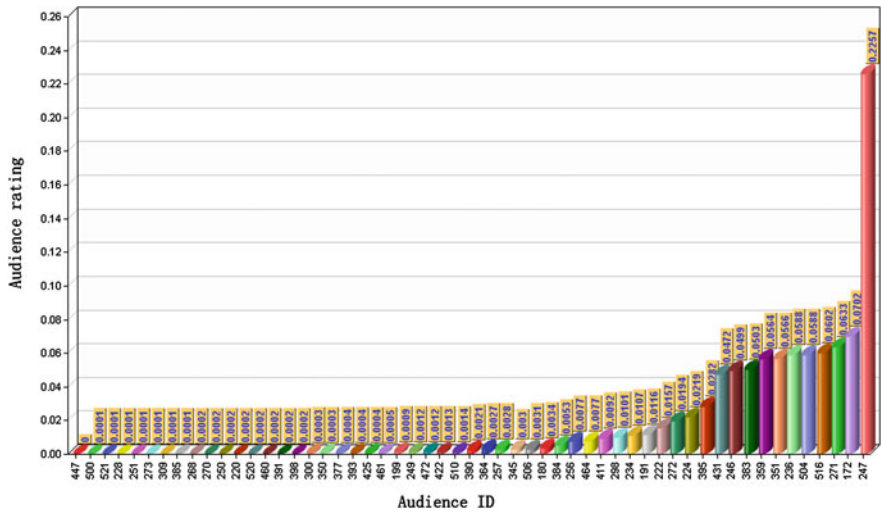
Fig. 5 The histogram of audience rating
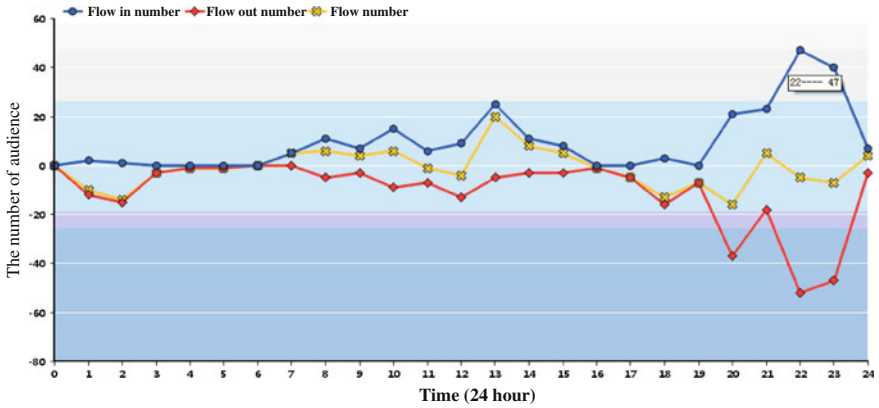
Fig. 6 The curve of audience flow

Fig. 7 The chat of audience crowd clustering
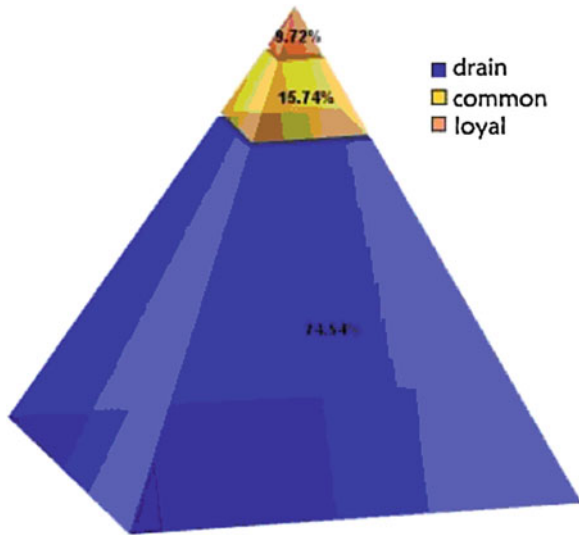
5 Display for Synthetic Broadcasting and Television DSS

Display is the visual expression for synthetic broadcasting and television DSS. Figures 5, 6 and 7 give the display of structured, half structured and non structured use case separately. Where, Fig. 5 gives the histogram of audience rating, Fig. 6 gives the curve of audience flow, and Fig. 7 gives the chat of audience crowd clustering.

6 Conclusion

The synthetic broadcasting and television DSS is practical and effective than traditional single data driven DSS or model driven model. It's combined the part of data base, model base and method base for quantitative analysis, the part of data base, model base and method base to reveal the essence of data, and the part of knowledge base and data mining for increasing the quantity and improving the quality of knowledge base. Then this paper analyzes the use case which divided into structured use case, half structured use case and non structured use case. The display of typical use case proves its validity.

References

1. Hersh MA (1999) Sustainable decision making: the role of decision support systems, systems, man, and cybernetics, Part C: applications and reviews. IEEE Trans 29(3):395–408
2. Chau VTN, Phung NH (2012) A knowledge-driven educational decision support system, computing and communication technologies, research, innovation, and vision for the future (RIVF). IEEE RIVF international conference, Ho Chi Minh City, pp 1–6
3. Wang S (2009) Application of decision support system in E-government, management and service science. Mass'09 international conference, Wuhan, China, pp 1–4
4. Yuan C, Zhang J, Yao J (2010) Present status and development of network convergence. J Beijing Univ Posts Telecommun 33(6):1–8
5. Schuster A, Sterritt R, Adamson K, Curran EP, Shapcott CM (2002) Towards a decision support system for automated testing of complex telecommunication networks, systems, man, and cybernetics, 2000 IEEE international conference, vol 3. Nashville 2002–2006
6. Faguo Z, Yang B, Li L, Chen Z (2008) Overview of the new types of intelligent decision support system, innovative computing information and control. 3rd international conference, Dalian. Liaoning, pp 267–270
7. Zhang M, Jiang Y, Wang M, Chunying Q (2011) Study on design of intelligent decision support system of engineering equipment repair, innovative computing information and control. International academic conference on numbers, intelligence, manufacturing technology and machinery automation, MAMT pp 36–41
8. Mentes A, Helvacioglu I, Hakki I (2012) Fuzzy decision support system for spread mooring system selection. Expert Sys Appl 3283–3297

Research on Breast Cancer Metastasis Related miRNAs Network Based on a Boolean Network

Wenying Zhao and Yafei Dong

Abstract Breast cancer (BC), which is one of the major malignant tumors, is a serious threat to women's health with noticeably increased mortality. BC starts as a local disease, but it can metastasize to the other organs. Metastatic BC is the major difficult factor for treatment of advanced breast cancer. It suggests that early diagnosis for patients is of vital importance. Recent studies proved the involvement of microRNAs (miRNAs) in BC metastases. It was found that miRNAs contribute to oncogenesis by functioning as down-regulated tumor suppressors, and also can function as over-expressed oncogenes. The goal of this study is to construct a comprehensive BC metastasis related miRNAs regulatory network by collecting and integrating relevant information, the analysis results based on a Boolean network shows that it has great biological significances and provides new perspective for the later research.

Keywords Breast cancer metastasis · Boolean network · miRNAs

1 Introduction

BC is the most common malignant disease in women. In these patients, it is not the primary tumor, but its metastases at distant sites that are the main cause of death, which is a serious threat to women's health [1]. Different levels of metastasis will

W. Zhao · Y. Dong
College of Life Science, Shaanxi Normal University, Xi'an 710062, Shaanxi,
People's Republic of China
e-mail: wenwen-805@163.com

Y. Dong (✉)
College of Computer Science, Shaanxi Normal University, Xi'an 710062, Shaanxi,
People's Republic of China
e-mail: dongyf@snnu.edu.cn

occur when BC develops to certain stage, which is the leading cause of death in patients. MiRNAs have been implicated in an increasing number of biological processes, including cancer. Several studies have now shown an involvement for these regulatory molecules in BC metastasis. The research on the relationship between miRNAs and BC metastasis may provide a new perspective to its detection and treatment.

MiRNAs are a recently discovered class of small regulatory RNAs, about 22 nt in length, that influence the stability and translational efficiency of target mRNAs. They are single-stranded RNAs that do not code for proteins themselves, but have a role in regulating the amount of protein expressed from coding RNAs. In particular, they can be used as gene network's input signal [2]. The analysis of miRNAs sequence shows that more than 50 % of the known miRNAs are located inside or close to fragile sites associated with cancer [3]. For example, the miRNA-17-92 is located at 13q31, a region commonly relate to lymphomas [4]. The involvement of miRNAs in the metastases process was initially found by Ma and his colleagues [5], that is, abnormal expression of miRNAs can cause steady state changes of gene regulation networks, which is the deep reason of the breast cancer metastasis.

2 Boolean Network

A Boolean network is a basic computational model but a powerful tool can describe gene network system. Since their inception in 1969 by Kauffman [6], Boolean network can provide insight into the overall behavior of genetic networks and is represented by variables with two possible states: on and off, denoted by 0 and 1 respectively, represented the individual genes states of the network. Each node takes a Boolean value of 0 or 1 representing the respective activation and inactivation status of the node. In Boolean network, interactions between genes can be represented by Boolean expression, where if regard each node as a gene, it can be seen as a gene network model. The properties that abstract the ultimate principle of complex system can lay the foundation for researching gene regulation networks by using Boolean network [7].

A Boolean network contains n elements (genes) $\{x_1,...x_n\}$. Each gene $x_i \in \{0, 1\}$ ($i = 1,..., n$) is a binary variable, the value of which at time $t + 1$ is completely determined by the values of some other genes at time t by means of a Boolean function f_i. For an n-node gene network, the state space is composition by state from $\{00... 0\}$ to $\{11... 1\}$, totally 2n states [8]. In this study, Boolean network model has been used to simulate a molecular pathway in BC metastasis related miRNAs network in order to understand the consequence of using logic gate rules between nodes.

3 MiRNAs Expression in BC Metastasis

Over the past few years, miRNAs that are aberrantly expressed in human breast cancer metastasis had been identification because of miRNAs profiling studies. Iorio [9] divided BC metastasis related miRNAs into two types: oncogenic miRNAs, which is up-regulated in tumor, denoted by 1; on the contrary, tumor suppressor miRNAs is down-regulated and denoted by 0. BC metastasis may be promoted by enhanced expression of miRNAs and down-regulation of miRNAs, and the imbalance between them.

This paper will focus on recent advances on the functions of miRNAs in BC metastasis, and the aberrant miRNAs expression is described in detail based on reported targets of BC metastasis associated miRNAs.

3.1 Oncogenic MiRNAs

MiR-10b was the first miRNA found to influence the metastatic capacity of cancer cells [5], which inhibits the translation of the transcription factor HOXD10 (homeobox protein D10), resulting in a cascade of cellular alterations that include expression of the prometastatic gene RHOC (Ras homologue gene family member C), a gene that promotes cancer cell migration and invasion.

Huang et al. [10] showed that miR-373 or miR-520c elicited a migratory phenotype in BC cell lines by collectively inhibiting expression of the metastasis repressor CD44, so that release the connection between cells and cell matrix and induce migration phenotype of BC cell lines.

The over-expressed miR-21 led the tissue inhibitor of metallopeptidase (TIMPs) down-regulation, which is known to be a target of miR-21 in breast cancer epithelial cell lines. Thus, miR-21 [11] increased the activity of MMP2 and MMP9 (matrix metallopeptidase). Extracellular matrix (ECM) degradation caused by increased MMPs is one of the key links in BC cells invasion and metastasis.

MiR-155 directly inhibited the expression of Ras homologue gene family member A (RhoA), a gene that regulates many cellular processes. The findings suggest that miR-155 is regulated by the TGF-β/Smad4 pathway and down-regulates RhoA protein expression to drive epithelial–mesenchymal transition (EMT) progression [12].

3.2 Tumor Suppressor MiRNAs

MiR-7 was found that consistently decreased in malignant tissue compared with normal breast tissue and its expression is correlates inversely with metastasis in human BC patients. MiR-7 was identified as a pivotal regulator of Focal adhesion kinase (FAK) in BC metastasis [13], which induces the EMT progression.

Table 1 Suppressor miRNAs and potential oncogenic miRNAs

Key miRNAs change in breast cancer metastasis			
Up-regulated (oncogen)	miRNA-21	Down-regulated (tumor suppressor)	miRNA-335
	miRNA-520c		miRNA-126
	miRNA-373		miRNA-146
	miRNA-155		miRNA-7
	miRNA-10b		miRNA-205

Tavazoie and his coworkers [14] found that abnormal expression of a series of miRNAs were closely related with and the metastasis of BC, in which miR-335 and miR-126 were identified as metastasis suppressor functioning through transcription factor SOX4 (SRY-related HMG-box) and one kind of ECM component, that is tenascin C (TNC). Kim et al. [15] furthermore identified miR-335 as a robust metastasis suppressor and tumor initiation suppressor locus in human breast cancer. In addition, another study [16] demonstrated that the loss of miR-146 expression promoted BC cell migration by maintaining level of Nuclear factor KB (NF-KB).

Down-regulation of E-cadherin is the crucial events in EMT and the key step of metastatic tumor at early stage. MiR-205 can interdict EMT process and tumor metastasis through inhibiting transcription repressors of E-cadherin, that is ZEB1 and ZEB2 (zinc finger E-box-binding homeobox).These studies suggest that miR-205 play important roles in Inhibiting tumor progression and metastasis [17].

The aforementioned abnormal expressed miRNAs in metastatic BC are summarized in Table 1 and the regulatory rules of this network are presented in Table 2.

The paper construct a regulatory network based on the above-mentioned rules so as to interpret how these miRNAs validated by biological experiment act as oncogenes and tumor suppressors in BC metastasis. As shown in Fig 1.

Table 2 Functions for the nodes in the BC metastasis related miRNAs network

Rules
miR-21—|TIMP—|MMP-9/MMP-2 → ECM degradation
miR-520c/miR-373—|CD44—|ECM degradation
miR-155—|RHOA—|EMT
miR-335—|TNC → ECM degradation
miR-126—|MMP-2 → ECM degradation
miR-146—|NF-KB → TWIST → EMT
miR-7—|FAK → EMT
miR-205—|ZEB1/ZEB2 → E-cadherin → EMT

Note Lines with *bars* indicate inhibition; *arrows* indicate promotion
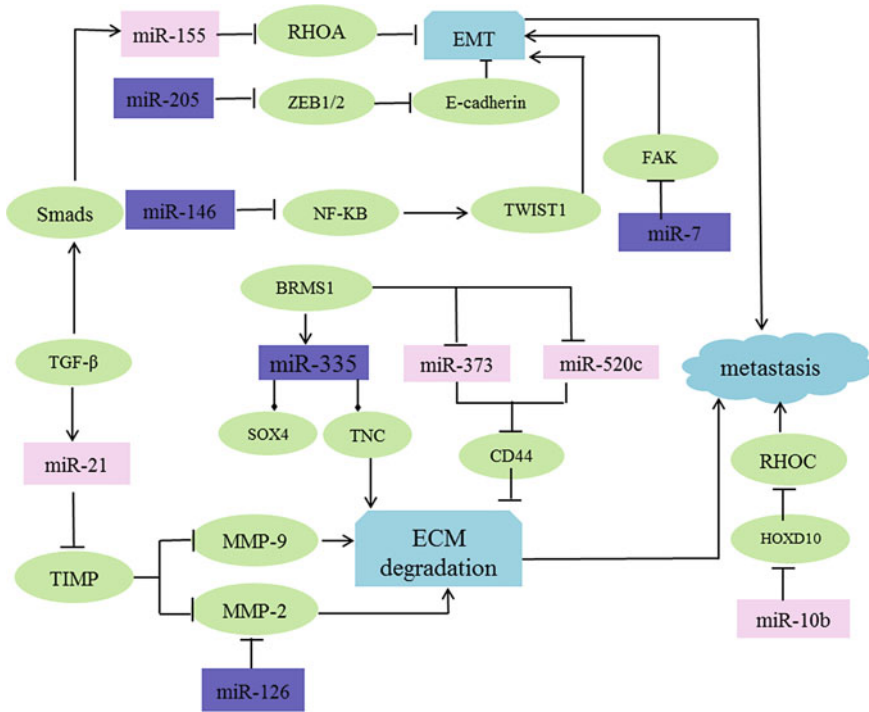
Fig. 1 Regulatory network of key miRNAs in BC metastasis. The miRNAs in *purple box* represent tumor suppressor and oncogenes are in *pink box*. *Lines* with *bars* indicate down-regulation, and *arrows* indicate up-regulation. Protein molecules and physiological effect are denoted by *light green ovals* and *rounded rectangles* respectively

4 Conclusion Analysis of Key miRNAs in BC Metastasis Regulatory Network Based on a Boolean Network

In order to focus on the input and output of system, which helps to develop a deeper understanding of the main functions of the network, complex biological functions were abstracted into space mapping relationship by {0, 1} and many of the details were ignored.

Take each component in miR-21 action pathway as the input (as show in Table 3), thus the output is 0 because of miR-21 is in low-level or loss of expression (denoted by 0) and have no effect on metastasis of cancer cells; On the contrary, the output is 1 when miR-21 is in high-level of expression (denoted by 1), that is, as a oncogenic gene, miR-21 plays a positive role in promoting metastasis and the result is coincided with the fact. As do the rest of the four oncogenic miRNAs (miR-520c, miR-373, miR-155 and miR-10b).

Table 3 Oncogenic miRNAs (e.g. miR-21)

Input				Output
miR-21	TIMP	MMP-9/MMP-2	ECM degradation	Metastasis
0	1	0	0	0
1	0	1	1	1
miR-21—|TIMP—|MMP-9/MMP-2 → ECM degradation				

Table 4 Tumor suppressor miRNAs (e.g. miR-205)

Input				Output
miR-205	ZEB1/ZEB2	E-cadherin	EMT	Metastasis
0	1	1	1	1
1	0	0	0	0
miR-205—|ZEB1/ZEB2 → E-cadherin → EMT				

Take each component in miR-205 action pathway as the input (as show in Table 4), thus the output is 1 because of miR-205 is in low-level or loss of expression (denoted by 0) and promoting metastasis of cancer cells; Conversely, the output is 0 when miR-205 is in high-level of expression (denoted by 1), that is, miR-205 inhibits the metastasis of cancer cells, which exactly consistent with this reality. The same as the rest of the four tumor suppressor miRNAs (miR-335, miR-126, miR-146 and miR-7).

Among the total ten miRNAs, in which miR-155, miR-205, miR-146 and miR-7 contribute to tumor suppressors or oncogenes of metastasis by functioning through EMT pathway. There are 2^4 cases all together (for the sake of brevity, not all of the truth value are shown here, but only some are available), in which miR-155 is overexpressed (denoted by 1) and the other three are in low-level or loss of expression (denoted by 0). EMT process is the initial stage of BC metastasis, which can be caused only when satisfies the above two given conditions. Thus, the output equals 1 and shows that BC cells are metastatic (shown in Table 5).

Table 5 MiRNAs functioning through EMT pathway

Input				Physiological effect	Output
miR-155	miR-205	miR-146	miR-7	EMT	Metastasis
0	0	0	0	0	0
0	1	0	0	0	0
0	0	1	0	0	0
⋮				⋮	⋮
1	0	0	0	1	1
⋮				⋮	⋮
1	1	1	1	0	0

Table 6 MiRNAs functioning through ECM degradation

Input					Physiological effect	Output
miR-21	miR-373	miR-520c	miR-335	miR-126	ECM degradation	Metastasis
0	0	0	0	0	0	0
0	1	0	0	0	0	0
0	0	1	0	0	0	0
⋮			⋮		⋮	⋮
1	1	1	0	0	1	1
⋮			⋮		⋮	⋮
1	1	1	1	1	0	0

As shown in Table 6, we can see miR-21, miR-373, miR-520c, miR-335 and miR-126 have effect on BC metastasis through ECM degradation pathway. There are 2^5 cases in all (same as above, some of the truth value were omitted only for concision), among which three miRNAs are in high-level expression simultaneously (denoted by 1) and the other two are in low-level or loss of expression (denoted by 0). ECM degradation, which promote the BC cells metastasis, can be initiated only satisfies the above two given conditions. Thus, the output equals 1.

5 Conclusion

This approach reveals that many protein genes involved in breast cancer are embedded in a miRNAs network. The paper briefly described the cancer network and discussed how miRNAs can act as oncogenes and tumor suppressors by a Boolean network which does not depend on parameters, a better understanding of the network and pathways regulated by these miRNAs will undoubtedly enable us to understand BC metastasis and therapy.

As to gene network research, people have begun to explore information processing ability of cells long ago. Shapiro [18] described a programmable finite automaton comprising molecules. Another simple molecular-scale programmable computer was also demonstrated in vitro by Shapiro's group [19]. Weiss and his colleagues [20] established a molecular computer in human kidney cells through RNAi technology, Xie [21] developed a synthetic regulatory circuit called "classifier", which can selectively identifies HeLa cells and triggers apoptosis without affecting normal cell types.

MiRNAs shows a great potential in tumor markers with the usage of new molecular technologies, such as DNA microarrays [22] and Next-generation sequencing [23]. Early research focused on traditional methods, such as immunohistochemical analysis [9], RT-PCR [24] and so on. Next, the research in this paper will focus on constructing a molecular computer about BC metastasis-associated miRNAs by using components such as promoters based on RNAi and

import the logic network into living cells through vector, allowing miRNAs as input data and biologically active molecules as outputs. More specifically, it can selectively identify whether BC cells with metastatic property. It is clear that the majority of this type of work remains to be done and it is likely that miRNA-based therapeutics will become a reality in the near future.

References

1. Toi M, Ohashi Y, Seow A (2010) The breast cancer working group presentation was divided into three sections: the epidemiology, pathology and treatment of breast cancer. Jpn J Clin Oncol 1:13–18
2. Garzon R, Calin GA, Croee CM (2009) MicroRNAs in cancer. Annu Rev Med 60:167–179
3. Calin GA, Sevignani C, Dumitru CD (2004) Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci USA 9:2999–3004
4. Ota A, Tagawa H, Karnan S (2004) Identification and characterization of a novel gene, C13 or f25, as a target for 13q31-q32 amplification in malignant lymphoma. Cancer Res 64:3087–3095
5. Ma L, Teruya-Feldstein J, Weinberg RA (2007) Tumor invasion and metastasis Initiated by microRNA-10b in breast cancer. Nature 449:682–688
6. Kauffman SA (1969) The large-scale structure and dynamics of gene control circuits. Ensemble Approach Theor 1:167–190
7. Kauffman SA (1993) The origins of order: self–organization and selection in evolution. Oxford University Press, New York, pp 460–462, 467, 479
8. Ranadip P, Ivan I, Aniruddha D, Michael L, Bittner ER (2005) Generating Boolean networks with a prescribed attractor structure. Bioinformatics 21:4021–4025
9. Iorio M, Ferracin M, Liu CG (2005) MicroRNA gene expression deregulation in human breast cancer. Cancer Res 65:7065–7070
10. Huang Q, Gumireddy K, Schrier M (2008) The microRNAs miR- 373 and miR-520c promote tumor invasion and metastasis. Nat Cell Biol 2:202–210
11. Milena S, Nicoloso S, Masayoshi S, Simona R, George AC (2009) MicroRNAs—the micro steering wheel of tumor metastases. Nat Rev Cancer 9:293–302
12. Kong W, Yang H, He L (2008) MicroRNA-155 is regulated by the transforming growth factor beta/Samd pathway and contributes to epithelial cell plasticity by targeting RhoA. Mol Cell Biol 22:6773–6784
13. Kong X, Li G, Yuan Y, He Y, Wu X, Zhang W, Wu Z, Chen T, Wu W, Lobie PE, Zhu T (2012) MicroRNA-7 inhibits epithelial–to–mesenchymal transition and metastasis of breast cancer cells via targeting FAK expression. PLoS ONE 7(8):e41523
14. Tavazoie SF, Alarcon C, Oskarsson T (2008) Endogenous human microRNAs that suppress breast cancer metastasis. Nature 7175:147–152
15. Png KJ, Yoshida M, Zhang XHF (2011) MicroRNA-335 inhibits tumor reinitiation and is silenced through genetic and epigenetic mechanisms in human breast cancer. Genes Dev 25:226–231
16. Hurst DR, Edmonds MD, Scott GK, Benz CC, Vaidya KS, Welch DR (2009) Breast cancer metastasis suppressor 1 up-regulates miR-146, which suppresses breast cancer metastasis. Cancer Res 4:79–83

17. Gregory PA, Bert AG, Paterson EL (2008) The mir-200 family and mir-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. Nat Cell Biol 5:593–601
18. Benenson Y, Paz–Elizur T, Adar R, Keinan E, Livneh Z (2001) Programmable and autonomous computing machine made of biomolecules. Nature 22:430–434
19. Benenson Y, Gil B, Ben–Dor U, Adar R, Shapiro E (2004) An autonomous molecular computer for logical control of gene expression. Nature 429:423–429
20. Rinaudo K, Bleris L, Maddamsetti R, Subramanian S, Weiss R, Benenson Y (2007) A universal RNAi–based logic evaluator that operates in mammalian cells. Nat Biotechnol 7:795–801
21. Xie Z, Wroblewska L, Prochazka L, Weiss R, Benenson Y (2011) Multi-input RNAi-based logic circuit for identification of specific cancer cells. Science 6047:1307–1311
22. Rajnish K, Anju S, Rajesh KT (2012) Application of microarray in breast cancer: an overview. J Pharm Bioallied Sci 1:21–26
23. Russnes HG, Nicholas N, James H (2011) Insight into the heterogeneity of breast cancer through next-generation sequencing. J Clin Invest 10:3810–3818
24. Shaomei F, Hongwei Z, Mofang L, Shuai J, Minghua L (2011) Expression of miR-155 and miR-21in breast cancer tissue. Chin J Exp Surg 8:1240–1241

Author Biographies

Wenying Zhao (1987), female, postgraduate of Shaanxi Normal University, her major research fields is gene network.

Yafei Dong is the corresponding author, his main research areas are molecular computing and gene network.

Design of Gas Monitoring System Based On Embedded ZigBee Module

Rongrong Gu, Kelei Sun, Wei Hao and Ping Ren

Abstract In order to solve the problems of traditional mine gas monitoring system, a design proposal based on wireless sensor network technology is put forward, which takes ARM as control core. The gas concentration data is collected by underground wireless sensor node, and converge on a central node by ZigBee network. At last, the data is sent to the monitoring center. Then the monitoring center analyzes and processes data received and send corresponding instructions to sensor network according to the results of data processing. On the basis of the above steps, real-time monitoring of mine gas concentration is achieved in this system. This system overcomes the shortcomings of the traditional mine gas monitoring methods, which has some merits such as flexible networks, low cost and strong reliability etc.

Keywords WSN · ZigBee · Gas monitoring system

1 Introduction

In recent years, more and more disastrous explosion accidents have happened in China's coal production, which were caused by gas accumulations and gas densities higher than specified levels. So gas disasters are still formidable enemies for safe production in coal mines, which also severely stop us from achieving highly efficient and high-yield work faces.

Currently, most of gas monitoring system in coal mines adopt wired network consisting of fixed sensors, which have the following disadvantages: because the structure of mine is complex, it is difficult to lay and maintain communication lines

R. Gu (✉) · K. Sun · W. Hao · P. Ren
School of Computer Science and Engineering, Anhui University of Science and Technology, Huainan 232001, China
e-mail: gurr@aust.edu.cn

and such lines are also easily damaged during workers' operations. Sensors are installed on fixed points, which cannot meet requirements of continuously dynamic of driving working faces. Monitoring points are relatively fixed, Therefore it easily result in monitoring blind area. Wireless sensor network has provided an effective technical means. It can form a multi-hop self-organizing network to make up for the weakness of wired devices. This article has proposed a design plan for a gas monitoring system based on embedded ZigBee.

2 Overall System Structure

This monitoring system is divided into signal acquisition part, signal transmission part and signal processing part. The overall structure of the system is as shown in Fig. 1.

Signal acquisition part is composed of wireless sensor nodes distributed over the monitoring area, which collects and pre-processes data information of gas concentration in the area and transmits the information through wireless communication modules. In signal transmission part, ZigBee central node converges all data from sensor nodes and sends the data to monitoring center through industrial Ethernet. Signal processing part consists of a host PC, which analyzes and processes data received and send corresponding instructions to sensor network according to the data processing results.
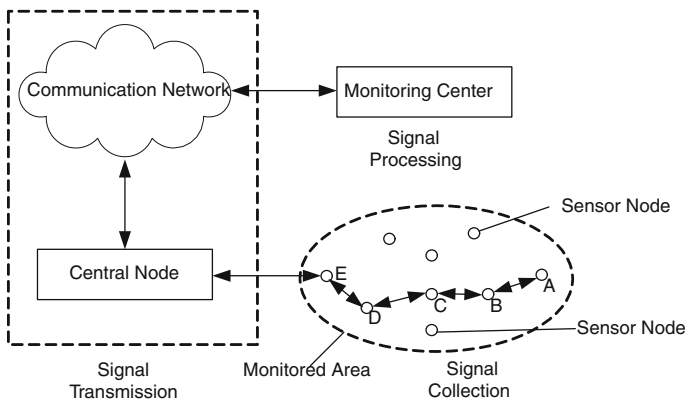

Fig. 1 The system structure diagram

3 Hardware Design of Sensor Node

Wireless sensor node of the system consists of four parts: a processor module, a sensor module, a wireless communication module and a power supply module [1], which are as shown in Fig. 2.

3.1 Processor Module

Processor module is responsible for the control of the whole sensor nodes operation, stores and processes data which acquired by itself and sent from other nodes. Core processor is a powerful and ultra-low-power 32-bit micro-processor LPC2294 launched by PHILIPS, which enjoys rich on-chip resources and can completely meet general industrial control needs. At the same time, it can also reduce design complexity of system hardware and enhance system stability as well as facilitate software and hardware upgrade in future. If wireless sensor network technology and control technology based on ARM microcontroller is combined, real-time gas monitoring system with high reliability and high stability can be realized. Functions of micro-processor are as shown in Fig. 3.
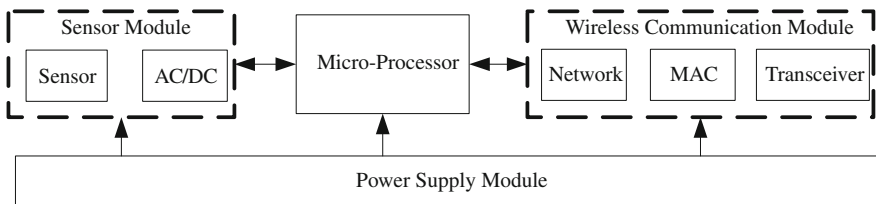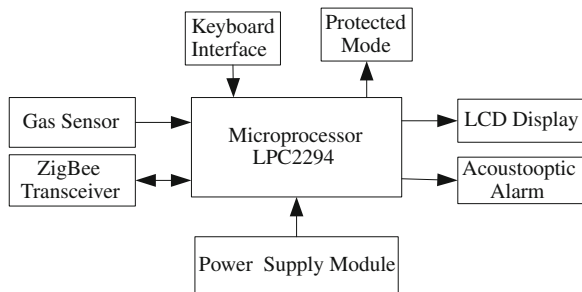

Fig. 2 The system structure diagram of sensor nodes [1]

Fig. 3 The function diagram of microprocessor

3.2 Sensor Module

Sensor module acquires and converts gas concentration data in monitored area. Gas sensor's working principle is that thermal catalytic element occur flameless combustion under the action of gas, which increases temperature of sensor components and changes resistance and voltage. Then gas concentration changes can be detected by measuring the voltage. LXK-6 catalytic element is adopted by the author, which is a kind of broad-spectrum gas sensitive element. It is simple in structure, high in sensitivity, sound in stability and low in power, and it is applicable for detecting mine gas concentration.

3.3 Wireless Communication Module

Wireless communication module is responsible for wireless communication with other sensor nodes, it means that these nodes exchange control messages with each other and send or receive the acquisition data wirelessly. This module is made of CC2430, combined with a high-performance 2.4 GHz RF transceiver core and an industrial compact and efficient 8051 processor. It has an 8 Kbyte RAM and a powerful peripheral module, so the design of peripheral circuit is simplified. The chip is very applicable to systems requiring ultra-low power. Functions of the wireless communication module are as shown in Fig. 4.

3.4 Power Supply Module

In wireless sensor network, power supply module provides the energy needed for sensor nodes, so selection of a good power supply module is crucial for the whole network. This module adopts two power supply modes—battery supply and external auxiliary supply. Battery supply can be adopted in case of unavailability
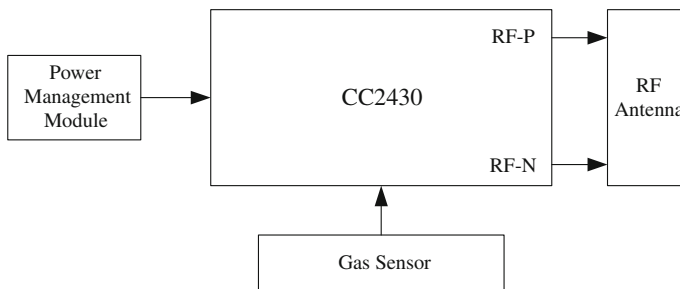
Fig. 4 The function diagram of wireless communication module

of external auxiliary power supply, and external power supply can be used if it is available. PAM2400 voltage stabilizing chip is adopted, which is low in power consumption and compact in packaging [2], It can stabilize voltage in 3.0 V, which will elongate service life of power supply.

4 Software Design of Sensor Node

4.1 Design of Node Communication

Major task of sensor nodes are to collect mine gas concentration and send the acquired data to ZigBee central node. To reduce energy consumption, sensor nodes will be sleep when no data need to be acquired or transmitted. They will be woken up once they need to perform normal work, and once again go to sleep after finishing the work. Flow chart for sensor node's working procedure is as shown in Fig. 5.
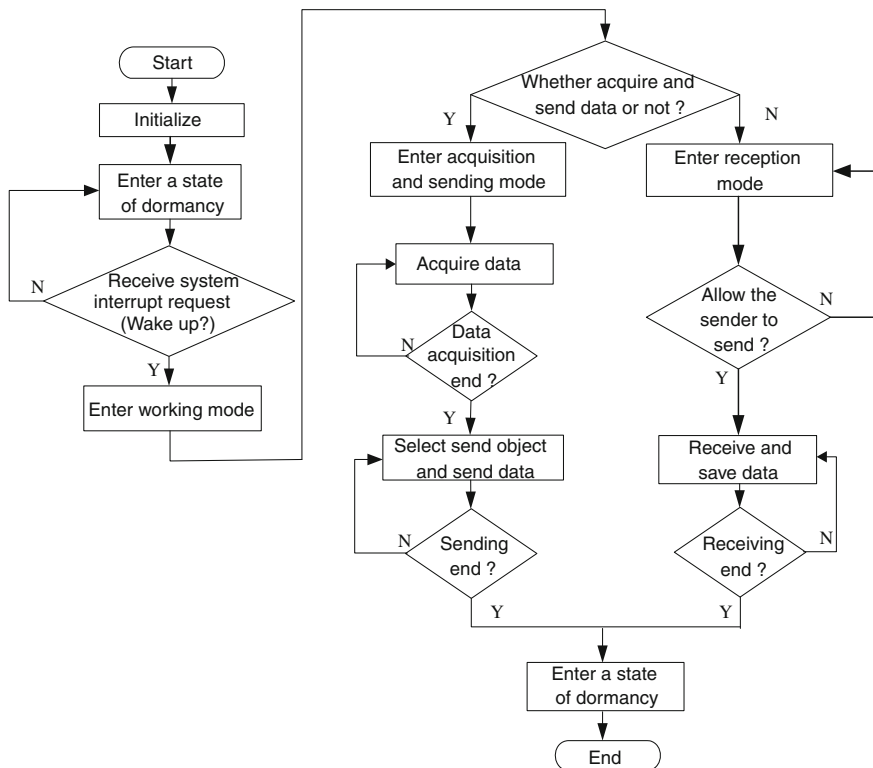
Fig. 5 The work flowchart of sensor node

4.2 Software Design of ZigBee Central Node

ZigBee central node need to builds the wireless network, manages sensor nodes and receives the data sent from sensor nodes. It connects wireless network with communication network and sends gas concentration information it has received to monitoring center. Its working flow is as shown in Fig. 6.

5 Conclusion

The author of this article has designed a gas monitoring system based on embedded ZigBee module. The system combines wireless sensor network with micro-processor technology based on ARM which will overcome shortcomings of current gas monitoring systems, such as huge investment on laying of communication lines, difficulties in line maintenance and existence of monitoring blind area. It can realize continuous, fast and real-time monitoring of gas concentration in coal mine. In addition, the system is characterized with networking flexibility,
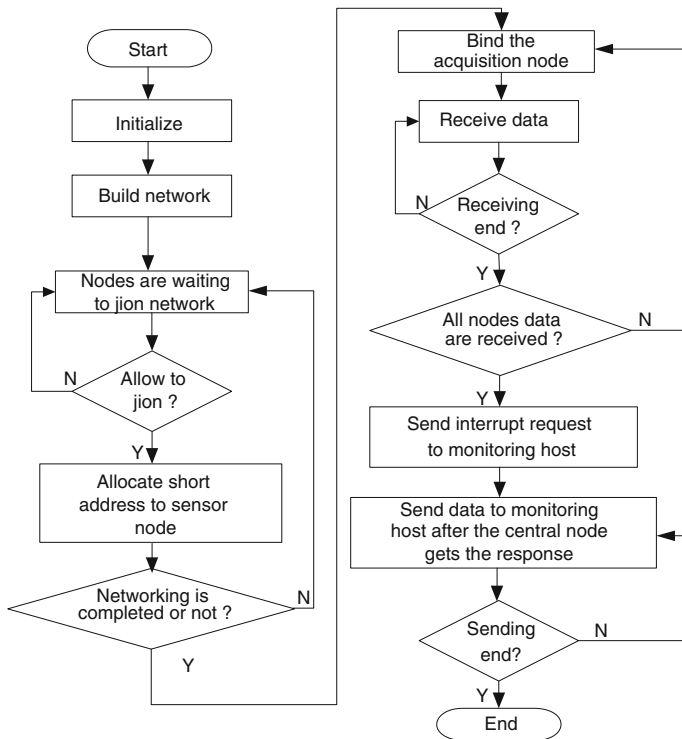
Fig. 6 The work flowchart of ZigBee center node

low cost, strong reliability and less system maintenance workload, so it can perfectly adapt to dynamic expansions of stope working face and provide a reliable guarantee for safe production of coal mine.

References

 1. Sun L, Li J, Chen Y, Zhu H (2011) Wireless sensor network. Tsinghua University Publishing House, Beijing
 2. Wu S, Huang Y, Shi M, Wang F (2012) Design of coal mine gas monitoring system based on wireless sensor networks. Coal Mine Mach 33(05):240–241
 3. Wang R, Sun L (2011) Wireless sensor network and its applications. Posts and Telecom press, Beijing
 4. Liu Y (2011) Introduction to internet of things. Science Press, Beijing
 5. Shen Y, Pei Q, Ma J, Pang L (2010) Introduction to wireless sensor network security technology. Posts and Telecom press, Beijing
 6. Shang F (2011) Communication protocols of wireless sensor network. Publishing House of Electronics Industry, Beijing
 7. Pan H, Dong Q, Zhang G, Yu L (2011) Operating system TinyOS of wireless sensor network. Tsinghua University Publishing House, Beijing
 8. Yuan H, Qi H, Ke X, Wang H, Han H (2011) ZigBee wireless sensor network for intelligent monitor system of pump. Comput Eng Des 32(2)
 9. Ding C, Zhao Y, Zhang M, Fu S (2012) System design of harmful gas information collection based on ZigBee. J Tianjin Polytech Univ 31(1)
10. Han Z, Wang L, Zhou Z (2012) Design of mine working face wireless sensor network monitoring system based on Zigbee. Coal Mine Mach 33(01):251–253
11. Chen S, Zhao T, Gao J (2011) Mine gas wireless monitor system based on ZigBee PRO. Coal Technol 30(09)

Berth-Crane Allocation Model and Algorithm of Container Berths in the Uncertain Environment

Jian-xin Liu, Yu-yue Du, Yong-fa Hong and Lin-lin Sun

Abstract According to the randomness of the vessel's arrival time and handling time, the establishment of a randomly-oriented environment container berths—crane allocation model, the optimizing goal is to minimize the average ship-waiting time. Taking into account the complexity of the model solution, this paper offers the design of an improved genetic algorithm to reduce the searching space, and according to the characteristics of the optimal solution. With testing example to verify that the model can simulate decision-making environment of berths—crane allocation problem and reflect the decision-maker's attitude toward risks and preferences. The algorithm can gain a stable and satisfactory solution within the operating time.

Keywords Berth crane allocation · Mathematical model · Genetic algorithm · Optimization

J. Liu (✉) · Y. Du · Y. Hong
College of Information Science Engineering, Shandong University of
Science and Technology, Qingdao 266590 Shandong, People's Rebpulic of China
e-mail: sdustljx@163.com

Y. Du
e-mail: yydu001@163.com

Y. Hong
e-mail: hzycfl@126.com

Y. Du
The Key Laboratory of Embedded System and Service Computing,
Ministry of Education, Tongji University, Shanghai 200092, People's Rebpulic of China

L. Sun
College of Civil Engineering and Architecture, SUST, Qingdao 266590 Shandong,
People's Rebpulic of China
e-mail: sunlinlin2003@126.com

1 Introduction

Container Terminal, an irreplaceable strategic position in the international logistics and national economy, is an integral part of the international logistics nodes. Since transport ships and port equipment are expensive, the transport ship's berthing span has a direct impact on economic efficiency and competitiveness of the port. Port-berths and quay-crane scheduling management is one of the important factors to determine the transport ship berthing span in the port. However, due to the constantly changing socio-economic environment in which the port is situated, and many random factors in production system, and diversity and uncertainty of customers' needs, more reasonable and effective planning on port berths and quay crane is needed to solve the problem, to improve the efficiency of port operations, to reduce costs and prices and eventually to attract more ships docked and more economic profits.

In recent years, the quay-crane allocation problem of the container terminal berth has been explored one after another. Lai and Shih [1] proposed a heuristic algorithm for the berth allocation problem. To minimize ship waiting time Imai [2, 3] etc. established objective nonlinear integer programming model to simulate the static, dynamic berth allocation (SBAP and DBAP). The above methods and models divided the coastline into several separate berths to be allocated. Lim [4] proposed an effective heuristic algorithm for solving continuous-shoreline berth allocation problem. Kim [5] employed mixed integer programming model to determine the ship's docked position and dwell time, with simulated quenching algorithm making the model approximate optimal solution. Li et al. [6] took berth—crane allocation problem as a scheduling problem to handle multiple tasks at the same time, and assumed that all ships waiting to harbor in Hong Kong, with minimum berth time as the target, and proposed heuristic algorithm. Similarly, Guan et al. [7] also took berth—crane allocation problem as a scheduling problem, and the optimization objective was to minimize the task completion time for crane allocation problem.

The above model and method is oriented at deterministic environment. But the actual widespread variety of uncertainties makes the deterministic model unable to truly reflect the actual problem, and to some extent they affect the allocation of decision-making, and even determine the final allocation scheme. Therefore, with undetermined variables to simulate uncertainties, this paper establishes container berths quay-crane allocation model for uncertain environment, and constructs heuristic algorithm to solve the model.

2 Description of the Problem

Under normal circumstances, before or after the container ship's arrival at the port, the container terminal scheduling staff, based on the Ship Harbor Program and other related information, quay-crane scheduling policy, offer shore handling

Fig. 1 Berth-quay planning
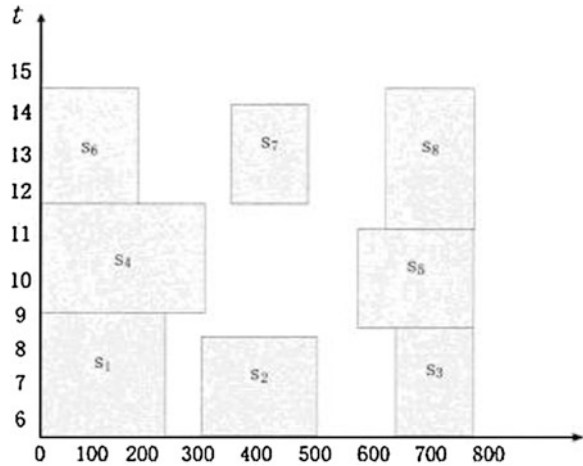
machinery assigned to the ship. Large overseas ship's scheduling timetable is given basically about a year in advance. Through EDI converted from the berth to the operation of the terminal, in addition to the ship and loading bridge (not all loading bridges are available), technical data, others like the captain and the length of the loading bridge arm must also be considered. All boats moored in a given time period should be mapped to the berth allocation system. Ship's arrival, berthing, loading and unloading operations and the departure are involved. Each ship should be allocated several cranes responsible for loading and unloading. In-harbor span is waiting time plus berthing-operations time. Optimize the allocation of the berth a crane through the appropriate allocation of berths, select the appropriate berthing order and equipped with a reasonable number of cranes to the ship in the port shortest, in order to optimize ship's turnaround utilization and meet customer's need, and reduce operating costs of the pier.

Like previously dealing with berths and quay crane, we use a graphical representation of the two-dimensional plane and ship rectangles, the X1 axis represents the length of the berths occupied; the X2 axis represents the number of quay crane and Y axis represents occupied length of time. So that one can get a clear, intuitive image of berth crane joint scheduling. Compared to the above separate two-dimensional one, berth-crane scheduling gets a new parameter, so the allocation is relatively complex; the graphical representation of the two dimensional plane schemes also gets relatively complex (Fig. 1).

3 The Berth Quay Planning Model

With the berth quay planning model and restrictions, mathematical equation is applied for the description and characterization of variants on the berth and quay crane.

3.1 Model Assumptions

Before establishing berth-crane model, it needs to make certain assumptions, and the assumptions involve two parts:

For the berth assumptions: (1) initial port berths, quay crane unoccupied; (2) The ship is based on a continuous form of berthing; (3) ships can get terminal services after arrival, the time of arrival as undetermined variables to consider; (4) the importance of each ship is determined according to the time requirements of the ship and the maximum acceptable waiting time; (5) loading and unloading time will be determined based on the location of berths, the number of quay crane and productivity and other factors; (6) berths to meet the ship's physical conditions, such as: water depth and shoreline length of the berth to meet the ship's requirements; (7) ships with only one-chance berthing, and no shifting berthing.

For the quay crane assumptions: (1) the crane can only serve one ship at a time; (2) a ship can be served by more than one quay crane; (3) the serving crane number is less than the ship's max allowable cranes: (4) The crane locations can not cross; (5) crane work has no effect on the efficiency of loading and unloading operations;

Model parameters Related concepts and symbols:

Tv, Tb, Tc, H to consider the number of ships calling at ports, the number of available berths and quay cranes, and time spans.

V, B, C, T, O one to consider the set of arriving ships, available berths set, time set, and the order set of the loading and unloading of ships respectively:

$V = \{1,2,\ldots, Tv\}$, $B = \{1,2,\ldots, Tb\}$, $C = \{1,2,\ldots, Tc\}$, $T = \{1,2,\ldots, H\}$, $O = \{1,2,\ldots, Tv\}$.

l_j :	The length of the ship j (including the level of security)
b_j :	Ship j berths location, b_j^0 :: ship j desired location, $\Delta b_j = \left| b_j^0 - b_j \right|$ :: the amount of deviation of the ship of j and hope berth
s_j :	Docked Ship at berth time, e_j :: for the ship's berth leaving time, $j \in V$:
m_j :	Completing hours of loading and unloading the ship j for the quay crane
r_j^{min} :	The ship j can be assigned to at least crane hours; r_j^{max} : can be assigned to the ship j max hours up to crane;
$R_j = \left[r_j^{min}, r_j^{max} \right]$:	shore the bridge-hour range can be assigned for the ship
vb_k :	K-loading and unloading ships assigned berth, $k \in O$.
ξ_{ik}^{tbs} :	The available time of the start of the service for the berth i of the k-th parked, $i \in B$:, $k \in O$.
EPT_j :	The scheduled time for the berthing of the ship j, $j \in V$.
LPT_j :	The latest time to leave the berth of the ship j, $j \in V$.
$r_{jtq} := 1$	indicates a ship j at time t is assigned q crane; otherwise $= 0$
c_j^1 :	Cost coefficient, waiting for berthing time for ship j

c_j^2 :	The time for ships j finished beyond the need to pay the cost factors
c_j^3 :	Hours for the ship j used in the containers crane cost factors
$r_{jt} = 1$:	if ship j at time t is arranged at least one quay crane, otherwise $= 0$
$u_j = 1$:	if ship j handling finished outside the time, otherwise $= 0$
$y_{jk} = 1$:	if the ship j berth when next propagation k, otherwise $= 0$
$Z_{jk} = 1$:	j loading and unloading of the ship after the end of the ship k began berths $= 0$ otherwise. wb_k: stands for k ship's waiting time before berthing.

3.2 Model Descriptions

Consecutive berth in the undetermined environment—crane allocation problem can be used as follows with undetermined variable parameter non-linear 0 1 programming model:

$$minZ = \sum_{j \in V} c_j^1 wb_k + c_j^2 u_j + c^3 \sum_{t \in T} \sum_{q \in C} (qr_{jtq}) \qquad (1)$$

$$\sum_{t \in T} \sum_{q \in R_i} (q^\alpha r_{itq}) \geq (1 + \beta \Delta b_i) m_i \quad \forall i \in V \qquad (2)$$

$$\sum_{i \in V} \sum_{q \in R_i} (qr_{itq}) \leq Q \quad \forall t \in T \qquad (3)$$

$$\sum_{q \in R_i} r_{itq} = r_{it} \quad \forall i \in V, \forall t \in T \qquad (4)$$

$$\sum_{t \in T} r_{it} = e_i - s_i \quad \forall i \in V \qquad (5)$$

$$(t + 1)r_{it} \leq e_i \quad \forall i \in V, \forall t \in T \qquad (6)$$

$$tr_{it} + H(1 - r_{it}) \geq s_i \quad \forall i \in V, \forall t \in T \qquad (7)$$

$$\sum_{j \in V} \sum_{q \in R_i} (qr_{itq}) \leq T_c \quad \forall j \in T \qquad (8)$$

$$p\left(\xi_{vb_k}^{tbs} - EPT_j < w_{bl}\right) \geq \phi \quad \forall i \in V \qquad (9)$$

$$Mu_i \geq e_i - LFT_i \quad \forall i \in V \qquad (10)$$

$$b_j + M(1 - y_{ij}) \geq b_i + l_i \quad \forall i,j \in V, i \neq j \qquad (11)$$

$$s_j + M\left(1 - z_{ij}\right) \geq e_i \quad \forall i, j \in V, i \neq j \tag{12}$$

$$y_{ij} + y_{ji} + z_{ij} + z_{ji} \geq 1 \quad \forall i, j \in V, i \neq j \tag{13}$$

$$s_i, e_i \in \{EST_i \ldots H\} \quad \forall i \in V \tag{14}$$

$$b_i \in \{0, 1, \ldots L - l_i\} \quad \forall i \in V \tag{15}$$

$$r_{itq}, r_{it}, u_i, y_{ij}, z_{ij} \in \{0, 1\} \quad \forall i, j \in V \; \forall t \in T \; \forall q \in R_i \tag{16}$$

Take (1) as the objective function, targeting that all vessels in port are required to spend minimum cost, which consists of three parts: the cost of waiting for berthing, the punitive costs of deferring and the minimum cost of loading and unloading time; (2)–(15) for the constraints with specific meanings as follows:

(2) Refers to the crane hours of loading and unloading under the circumstances of crane interference or deviation from the expected berthing location, with α as the crane interference index, β as deviation from the berth factor.
(3) Represents the number of quay crane to be assigned to be assigned to ship j.
(4) is the consistency of variables r_{jtq} and variable r_{jt} assignment, ship j at time t, certainly can guarantee at least a quay on service.
(5) The ship's service time is the ship j end time minus the start time.
(6) All ships have one and only one berth opportunity.
(7) Ensures that no more than one ship is moored in the berth at a given time.
(8) Requires that the ship berthing waiting time has the guaranteed rate ϕ.
(9) The completion time of the ship's service is less than the latest finish time.
(10) The length of the ship set the variables to consider.
(11) The ships' docking order set the variable Z_{jk}.
(12) Time and space constraints, avoid overlapping on the space–time diagram.
(13)–(16) are the constraints of the value of the variable.

4 Solution of the Model

4.1 Chance Constraints into Deterministic Equivalence Class

Chance Constrained programming is an important branch of stochastic programming, first proposed by charnes and Cooper, focusing on containing random variable constraints, and decision-making must be observed before the realization of random variables. In view that the decisions made in the unfavorable situation may not satisfy the constraint conditions, a principle can be adopted: decisions can be made with the constraint conditions unmet to a certain extent, but the deviation of the constraint condition is established is not beyond a certain confidence level.

Chance constrained transport planning model and deterministic planning model are distinct in that the former involves chance constraints and then this paper detects the respective equivalent forms and transforms it into a deterministic model solution.

4.2 Heuristic Algorithms

Given the complexity of the model solution, the paper adopts heuristic algorithm to solve the model. Ships to berth are in accordance with the order of berths. When a new ship j berths, Insert (j) operation, the ship j to be inserted into the berth plan partial order, assigned to a ship j berth time sj, berth position bj, the need to deploy the number of quay crane q, Insert (j) operation as shown in Fig. 2:

Explanation of each step is given below:

1. Z represents a cost of parked ship j, and its initial value is set to infinity.
2. Set the ship j berthing time for the scheduled berthing time of the ship j.
3. Set the parked position to set the ship j's expected location b_j^0
4. Allocate cranes to the moored ship j according to berth allocation in the space–time diagram, following allocation principle of optimal crane processing time of the ship j, reduce the berthing time of the ship j.
5. Suppose a cranes is assigned to the ship j, then the processing time of this ship is to definite, then it needs to check whether there is overlap between the ships in the berth space–time diagram.
6. If there is no overlap, the berthing cost of the ship j can be calculated, a new optimal solution can be obtained, i.e. in the space–time diagram, $\left(s_j^*, b_j^*\right)$ and
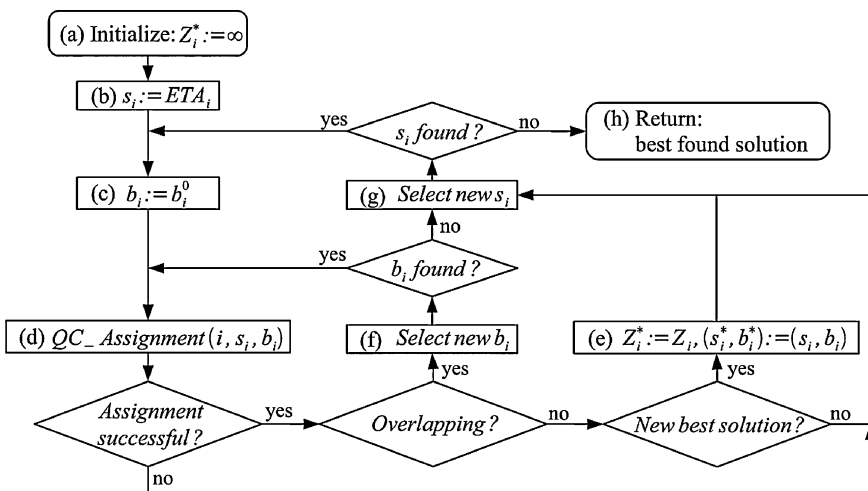

Fig. 2 Procedure insert(i) of the constructor

costs Z_j^* can be obtained. Other with overlapping berthing inappropriate for ship j and another park location should be chosen.

7. If can find a new position, return to the 4 with the re-allocation of quay crane.
8. In the following two situations it needs to choose a new berth time: If you do not find a new location or do not get a successful crane allocation.
9. If not find the new berth start time, and find the optimal solution $\left(s_j^*, b_j^*\right)$, insert (j) completion; with new berths start time, return and re-solve.

5 Summary

As for the quay crane allocation problem in container terminal berths, the arrival time of the ship, loading and unloading span and its subordinate time variables are all dynamic and uncertain. In this paper undetermined variables are applied to represent these variables, proposes undetermined-variable approach, based on chance constrained undetermined planning, establishes a container berth-crane allocation model. The tests shows that the container berth-crane distribution undetermined planning model can truly reflect the uncertainties in the above-mentioned issues, reflecting the risk attitudes and preferences of decision makers. Such could be further studied in the future as the uncertainty of arrival time scheduling uncertainties.

References

1. Lai KK, Shhi K (1992) A study of container berth allocation. J Adv Transp 26(l):45–60
2. Imai A, Nagaiwa K, Tat CW (1997) Efficient planning of berth allocation for container terminals in Asia. J Adv Transp 31(1):75–94
3. Imai A, Nishimura E, Papadimitriou S (2001) The dynamic berth allocation problem for a container port. Transp Res Part B 35:401–417
4. Lim A (1998) The berth scheduling problem. Oper Res Lett 22:105–110
5. Kim KH, Moon KC (2003) Berth scheduling by simulated annealing. Transp Res Part B 37:541–560
6. Li CL, Cai X, Lee CY (1998) Scheduling with multiple-job-on-one-processor pattern. IIE Trans 30:433–445
7. Guan YP, Xiao WQ, Cheung RK et al (2002) A multiprocessor task scheduling model for berth allocation: heuristic and worst case analysis. Oper Res Lett 30:343–350

Numerical Research on Coal Mine Ground Stress Field Based on Multi-Objective Optimization Method

Qinjie Liu, Xinzhu Hua and Ke Yang

Abstract With the depth and intensity of coal mining increasing, in situ stress plays more and more important roles in the surrounding rock control. This paper establishes a theoretical model for the in situ stress inversion problem, making use of multi-objective optimization method, and proposes an integrated stress analysis method, in which geo-mechanical method, finite element method, computer technology are adopted. This study method can acquire the magnitudes and directions of coalmine ground stresses via some discrete measured data, considering the basic influencing factors such as the folds and faults. One coalfield as the example of Huainan coalmine has been studied, and the results indicate that the relative errors are less than 10 %, which can mostly meet the requirement of the mining application. Then, the distribution of horizontal stresses has been predicted, which can provide an important reference for the stability design of mine openings, as well as the type and amount of ground support needed to maintain a safe working environment for coal miners.

Keywords Multi-objective optimization · Numerical research · Finite element method · Stress inversion

Q. Liu (✉) · X. Hua · K. Yang
School of Mining and Safety Engineering, Anhui University of Science and Technology, Huainan 232001, People's Republic of China
e-mail: qjliu@aust.edu.cn

X. Hua · K. Yang
Key Laboratory of Integrated Coal Exploitation and Gas Extraction, Huainan 232001 Anhui, People's Republic of China

1 Introduction

The knowledge of in situ stress state is indispensable for design and construction of coal mining engineering [1]. Both the magnitudes and directions of in situ stress in a coalfield have an important position for the excavation and support of the roadways and chamber in rocks [2]. This paper introduces a novel method for the analysis of stress field based on the multi-objective optimization method. The authors start with documenting the theoretical model of the ground stress field inversion, making use of multi-objective method, derive the theoretical and finite element models from the fundamental equations of elastic mechanics and the genesis of tectonic stress, and then establishes a basic investigation method that could deal with the problem of tectonic stress inversion for coal mining. The comparison between the studying results and tested data by over-coring method demonstrates that the precision of the research can be accepted by the experts on site.

2 Theoretical Model of the In-situ Stress

In the exploration and development of coal and coalbed methane, all the in situ stress information got by the researchers are only a few discrete data sets from a certain in-site tests. Therefore, we have to solve the inverse problems to get the stress field of a coalmine, which is regarded as the inversion analysis of differential equations in mathematics [3]. It is commonly perplexing to solve such inverse problems, and it seems to be impossible to get their analytical solutions, so the numerical calculating method is inevitable [4].

2.1 Multi-objective Problem and its Solving Methods

The optimization technique is quite a long-standing problem, and its solving methods are numerous. Sometimes there may be more than one designing objectives or parameters should achieve optimum values simultaneously, which is always named as multi-objective optimization problem. The typical form of the mathematical modeling for multi-objective optimization problem can be written as follows [5–7]:

$$
\begin{cases}
\min & f_1(\mathbf{X}) \quad \mathbf{X} = (x_1, x_2, x_3, \ldots, x_n)^T \in E^n \\
\min & f_2(\mathbf{X}) \\
\vdots & \vdots \qquad\qquad\qquad\qquad\qquad \vdots \\
\min & f_q(\mathbf{X}) \\
S.t & g_i(X) \leq 0 \qquad i = 1, 2, \ldots, m_1 \\
& h_j(X) = 0 \qquad j = 1, 2, \ldots, m_2
\end{cases}
\tag{1}
$$

It is often difficult to make all the constrained objectives reach the optimum values in the multi-objective optimization problems. Thus, the decision-making plays a significant role in this kind of multi-objective optimization problems, which has been proposed by Kuhn and Tucker in 1950s. Thereafter, many solving methods have been developed, such as analysis and decision method, globalization criteria method, hierarchy optimization method and inheritance arithmetic, etc.

The author has made a full comparison among all the above-mentioned methods, and finally picked out the hierarchy optimization method as the fundamental approach in this study [8]. The main idea of this method is as follows:

Step 1. Minimize the first objective function $f_1(\mathbf{X})$, and obtain the optimum value f_1^*, which can be written as:

$$
\begin{cases}
f_1^* = \min f_1(\mathbf{X}) & \mathbf{X} = (x_1, x_2, x_3, \ldots, x_n)^T \in E^n \\
S.t. \quad g_i(\mathbf{X}) \leq 0 & i = 1, 2, \ldots, m_1 \\
h_j(\mathbf{X}) = 0 & j = 1, 2, \ldots, m_2
\end{cases}
\tag{2}
$$

Step 2. Translate the first objective function into another restricting condition $w_1(\mathbf{X}) = f_1(\mathbf{X}) - f_1^* \leq 0$, and then minimize the second objective function $f_2(\mathbf{X})$, and attain the optimum value:

$$
\begin{cases}
f_2^* = \min f_2(\mathbf{X}) & \mathbf{X} = (x_1, x_2, x_3, \ldots, x_n)^T \in E^n \\
S.t. \quad g_i(X) \leq 0 & i = 1, 2, \ldots, m_1 \\
h_j(\mathbf{X}) = 0 & j = 1, 2, \ldots, m_2 \\
w_1(\mathbf{X}) \leq 0
\end{cases}
\tag{3}
$$

Step 3. In the same way, minimize and translate the first $(q-1)$ objective functions into restricting conditions step by step, and eventually optimize the last objective function, whose mathematical model is:

$$
\begin{cases}
f_q^* = \min f_{q-1}(\mathbf{X}) & \mathbf{X} = (x_1, x_2, x_3, \ldots, x_n)^T \in E^n \\
S.t. \quad g_i(\mathbf{X}) \leq 0 & i = 1, 2, \ldots, m_1 \\
h_j(\mathbf{X}) = 0 & j = 1, 2, \ldots, m_2 \\
w_k(\mathbf{X}) \leq 0 & k = 1, 2, \ldots, q\ 1
\end{cases}
\tag{4}
$$

There may be a phenomenon that the current objective function has no solution when all the previous ones have reached the optimum solutions. Thus it can be settled by relaxing the requirements for each solution, that is:

$$
w_j(\mathbf{X}) = F_j(\mathbf{X}) \left(f_j^* + f_j \right) \leq 0 \quad (j = 1, 2, \ldots, q\ 1)
\tag{5}
$$

Finally, the constrained problems can be transformed into the unconstrained ones, which are accomplished by means of penalty functions, and led to the following sub-problem statement.

$$
Q(x, q) = Minimize \left(\frac{F_1}{F_0} + \sum_{i=1}^{m} P_x(x_i) + q \left(\sum_{i=1}^{m_1} P_g(g_i) + \sum_{i=1}^{m_2} P_h(h_i) \right) \right)
\tag{6}
$$

Where Q is the unconstrained objective function, P_x, P_g, P_h are penalty function applied to enforce the constrained design and state variables, q is a penalty factor, and F_0 is a reference objective function that is selected from the current group of design sets.

2.2 The Inversion Model of Coal Mine Ground Stress

Suppose there is a geological space domain labeled V, including a certain faults and folds whose border can be labeled S. According to the FEM equation of three-dimensional boundary values problem, the relationship of the nodal stress of this geological model and its boundary loads can be expressed as [9]:

$$\sigma = DLK^1F(X) \tag{7}$$

Where σ is the nodal stress, $F(X)$ is the matrix of boundary loads, X is a parameter vector used for describing the unknown boundary loads endured by the geological space domain, D is an elastic matrix, L is the derivative operator matrix, and K is the globe stiffness matrix.

As is seen in Eq. (7), once the boundary loads matrix $F(X)$ is obtained, the stress field can be acquired easily with the auxiliary conditions of mechanical parameters and the boundary conditions. However, the boundary loads are always unknown, so it is prerequisite to settle the differential equation inversion problems making use of the above multi-objective optimization method.

Solving the coal mine stress field inversion problem is to find out the optimum model parameters or boundary conditions that could fit the tested data best under a certain criteria. In the procedure of our research, all the observation data are usually the stress magnitude and/or azimuth of some test points by the overcoring method using the CSIRO [10–12]. Select the resultant function of magnitude and azimuth of tested stresses as the optimizing objectives, calculate the stress by means of the finite element numerical method from the predefined parameters and boundary conditions, compare the calculated results with measured ones, seek and rectify the design variables utilizing the multi-objective optimization method until all the optimizing objectives achieve the minimum values. Thus, we can obtain the optimum boundary load parameters, then incorporate them into the FEM model, and finally access the ground stress distribution for coal mining.

3 Implementation of Coal Mine Ground Stress

We've designed a FEM analysis program to optimize the boundary loads of coalmine, applying the above-mentioned methods and techniques. Take one coalfield in Huainan (Anhui Province, China) as an example to illustrate the basic

routines and results of the coalmine ground stress-analyzing model based on multi-objective optimization method.

3.1 Establish the Geological Model

The authors established the simulation model based on the geological structure map of the study Mine Field, set the eastward direction as positive X-axes, and the northward direction as positive Y-axes, and then set the calculation domain to 6,000 × 2,500 m. For the sake of the calculation accuracy, most of the faults and other geological structures are included in the simulation model (Fig. 1).

It is commonly to determine the rock mechanical parameters based on the lab mechanics experiments and the existing research results about the surrounding areas. Recently, most of the researchers give the identical rock physical parameters to one studying stratum from the general macroscopic simulation effects, after the comprehensive consideration of the lithofacies changes. Treat the fault as the rupture zone, and designate a less modeling parameter in accordance with the fault scale.

3.2 Back Analysis of the Boundary Conditions

Since the targeted coaly stratum is relatively thin compared to its depth, the variation of the vertical component of tectonic stress can be neglected and the boundary loads in the horizontal plain hardly vary. Therefore, it is reasonable to optimize the coalmine ground stress field using the plane stress model.

1. The boundary conditions: (a) Hinge the bottom-left corner and the bottom-right corner of the model to avoid the rigid displacement; (b) Place the four border with uniform compressive pressure, labeled P_1, P_2, P_3, P_4, and the predefined initial values are 15 MPa; (c) Impose four uniform shear stress P_5, P_6, P_7, P_8, to each border, and give them 5 MPa as the initial values (see Fig. 1).

Fig. 1 Sketch map of the simulation model
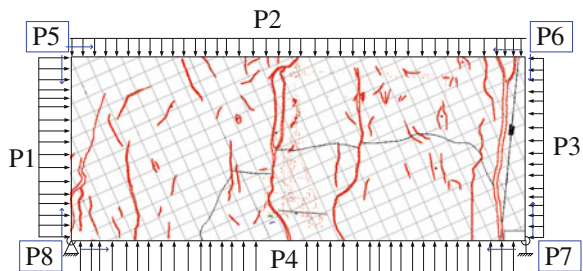
2. Objective function and state variables: Considering the in situ tested data and the lab experimental result, select the compound function of the magnitudes and direction angles of the horizontal principal stresses as the optimizing objectives, and set 9 values of the horizontal principal stresses as the state variables, which are compared and constrained to the calculated values in the simulation procedure.
3. Calculating conditions: Set the lower bound and upper bound of the border compressive pressure to 10 and 30 MPa. Designate the two limits of the shear stress are -5 and 5 Mpa respectively. In addition, the tolerance of the objective functions (i.e. τ) is set to 10^{-6}, namely, the iteration calculation will not end until τ achieves 10^{-6}.

3.3 Results of the Stress Inversion

The optimum boundary loads corresponding to all the items of 9 state variables have been acquired after the analysis and back-analysis, which can be listed in Table 1.

Table 2 shows the comparison of the coal mine ground stress simulation values under the above optimum boundary loads and the tested values, which illustrates that most calculating results for the magnitude and azimuth of ground stresses are approximately or quite close to the measured ones. Table 2 makes us come to conclude that the precision of the technique can meet the requirement of actual mine engineering.

Table 1 The optimum boundary loads from the coalmine ground stress inversion

Boundary loads	P1	P2	P3	P4	P5	P6	P7	P8
Values (MPa)	28.102	12.860	26.078	17.285	3.667	-2.795	4.902	-1.948

Table 2 Comparison between simulation results and tested values

Test point no.	Items	Tested values	Simulated values	Errors
KDZ-1	σ_H	30.15 MPa	28.53 MPa	5.37 %
	σ_h	21.05 MPa	19.72 MPa	6.32 %
	Azimuth	NE112.54°	NE115.5°	2.96°
KDZ-2	σ_H	31.97 MPa	33.85 MPa	5.88 %
	σ_h	14.04 MPa	15.60 MPa	11.11 %
	Azimuth	NE120.71°	NE116.45°	4.26°
KDZ-3	σ_H	37.27 MPa	34.52 MPa	7.38 %
	σ_h	20.63 MPa	22.78 MPa	10.42 %
	Azimuth	NE122.29°	NE120.15°	2.14°

Notes σ_H: maximum stress, σ_h: minimum stress, Azimuth: azimuth of σ_H
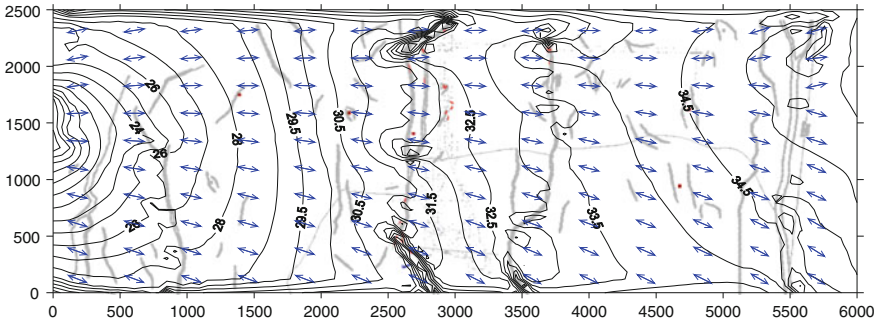
Fig. 2 The contour and azimuth map of the horizontal maximum stresses

4 Ground Stress Distribution of Study Coal Mine

Based on the previous theoretical method and calculation model, we can obtain the stress distribution of the study coalmine by analyzing and processing the calculating results with the special software programmed by the authors and their research group.

Figure 2 is the contour and azimuth map of the horizontal maximum stresses of the study coalmine, from which we can obtain the stress magnitudes and directions of the horizontal stresses everywhere the excavation or mining occurs. Figure 2 also shows that the horizontal maximum principal stresses mainly distribute between −24 MPa and −37 Mpa (therein minus suggests compressive stress), and the regional stress values increase gradually from west to east with the increasing depth of the coaly stratum. The azimuth of the maximum stress varies mainly from *NE*85° to *NE*127°, similar to the far-field stress direction in this area. It is clearly that the fold and fault formation plays an important role in the distribution of stress field, which mainly exhibits the following characters: (a) the stress magnitudes are seriously impacted by them, and the stress gradient near the fault zone is much higher than other parts; (b) the azimuth angles of the maximum stresses have deflected and dispersed clearly, especially those close to the complex-distributed faults regions.

5 Conclusions

1. A practical and effective method for back analyzing the in situ stress field of coal mine has been provided. The integration of the multi-objective optimization technique and the finite element method make it possible to acquire the boundary loads of coalmine ground stress simulation model, which can be used to investigate the ground stress field of the whole coalfield.

2. The stress optimizing inversion model has introduced multi-objective optimization technique and FEM to the process of the stress field analysis of coalmine, whose objective functions can integrate well with calculated stresses and in situ tested ones. In the pursuit of the coalmine ground stress analysis, one coalfield has been studied as the example, whose results demonstrate that the relative Errors between calculating results and measured conditions are mainly less than 10 %, which can meet the requirement of actual engineering application.

References

1. Cai M, Peng H (2011) Advance of in situ stress measurement in China. J Rock Mech Geotech Eng 3(4):373–384
2. Wu J, Cai D, Tong S (2009) Distribution characteristics of earth stress and its effect on tunnel deformation in Renlou coal mine. J Hefei Univ Technol (in Chin) 32(10):1554–1557
3. Wang R (1999) Earth's tectonic stress inversion, WANG Ren Corpus. Peking University Press (in Chinese), Beijin
4. Xu W, Liu S (2002) Stochastic back analysis of initial stress field in the complex rock mass and genetic optimization. In: Balkema AA (ed) Proceedings of Eurock, Madeira, Portugal (in Chinese)
5. Klimentos T, Ghosh A, Sagar R (2008) Cased hole sonic technique for stress determination and optimization of oriented hydraulic fracturing operations, SPE 113573
6. Tang H, Qin X (2000) Practical optimization method. Dalian University of Technology Press, Dalian
7. Zhang LQ, Yue ZQ, Yang ZF (2006) A displacement-based back-analysis method for rock mass modulus and horizontal in situ stress in tunneling illustrated with a case study. Int J Tunn Undergr Space 21:636–649
8. Liu Q, Yan X, Yang X (2009) Application of optimization back-analysis method in reservoir stress and fracture study. Pet Drill Tech (in Chin) 37(2):27–31
9. Yan X, Liu Q, Yang X (2009) Application of dislocation model in fracture prediction of low-permeability reservoir. Acta Pet Sin (in Chin) 30(2):252–258
10. Shen B, King A, Guo H (2008) Displacement, stress and seismicity in roadway roofs during mining-induced failure. Int J Rock Mech Min Sci 45:672–688
11. Zahl E, Dunford J, Larson M (2002) Stress measurements for safety decisions in longwall coal mines. In: 21st international conference on ground control in mining, Morgantown, 6–8 August, pp. 45–52
12. Guo Q, Ji D (2012) Study on measuring and test technology of ground stress field in No. 10 mine of Pingdingshan coal mining group. Coal Sci Technol (in Chin) 40(2):12–14

Mapping Loops onto Coarse-Grained Reconfigurable Array Using Genetic Algorithm

Li Zhou, Dongpei Liu, Min Tang and Hengzhu Liu

Abstract Coarse-grained reconfigurable array (CGRA) is a competitive hardware platform for computation intensive tasks in many application domains. The performance of CGRA heavily depends on the mapping algorithm which exploits different level of parallelisms. Unfortunately, the mapping problem on CGRA is proved to be NP-complete. In this paper, we propose a genetic based modulo scheduling algorithm to map application kernels onto CGRA. An efficient routing heuristic is also presented to reduce the mapping time. Experiment result shows our algorithm outperforms other heuristic algorithms both in solution's quality and mapping time.

Keywords Genetic algorithm · Loop mapping · CGRA

1 Introduction

The coarse-grained reconfigurable array is one of the best candidate architecture for digital signal processing domain such as wireless communication and image processing [1]. The high flexibility and efficiency of CGRA makes it a competitor of digital signal processor (DSP) and field programmable logic array (FPGA). Usually, the CGRA contains an array of processing elements (PE) which is controlled by context word in registers. The PEs can be reconfigured to perform different kinds of operations. However, the reconfigure is word level rather than bit level in FPGA, so CGRA is more efficient in running applications. Due to its large amount of processing resources, the instruction level and the loop level parallelism

L. Zhou (✉) · D. Liu · M. Tang · H. Liu
Institute of Microelectronics and Microprocessor, School of Computer Science,
National University of Defense Technology, Changsha, China
e-mail: zhouli06@nudt.edu.cn

can be exploited to improve performance, which makes CGRA more attractive in execution computation intensive tasks.

Many researches have been done on how to map an application kernel onto CGRA, most of them concern mapping loop kernels by module scheduling. The loop kernel is organized in a special order so that loop body can overlap each other with a start initial interval (II) delay. Since it has been proved NP-complete [2], several methods focus on efficient heuristic design. Mei [3] was the first to introduce modulo scheduling into CGRA mapping. This approach exploits loop level parallelism through iterative modulo scheduling. The loop pipelining technique in very large instruction word (VLIW) processors is applied through simulated annealing. Park [4] implements the procedure by utilizing edge-centric modulo scheduling. This method is route-aware and focuses on how to assign PEs according to edges rather than on the placement of nodes. However, these heuristics either have a very long running time or relatively inferior result.

In this paper, we propose a loop mapping algorithm based on the genetic algorithm (GA). The algorithm is a module schedule method that reduces the II of a loop kernel on a given CGRA. Near optimal solutions can be found within a reasonable time.

The rest of this paper is organized as follows. Section 2 defines modulo scheduling problem on CGRA. Section 3 shows the detailed algorithm. Section 4 presents the evaluations, and Sect. 5 concludes the paper.

2 Problem Definition

The loop level parallelism is usually performed by loop pipelining. It starts loop body before last iteration finished. An initial interval is required between two adjacent iterations because of the hardware resource requirements and data dependency constraints in loop body. The goal of module scheduling is to find an order for operations under resource and data dependency constraints so that II is minimized. Modulo scheduling for CGRA is much complicated than that on VLIW processor. Both of the placement and route problems need to be considered.

A loop kernel can be represented by a data flow graph (DFG) $G = \langle V, E \rangle$, where V is the operations in loop and E is the data dependency between operations. Each $e = \langle u, v \rangle \in E$ has a distance attribute $e.dif$. The $e.dif > 0$ indicates data consumer v need the result of producer u in the last $e.dif$ iteration. Given the target CGRA architecture, we also use a directed graph $C = \langle P, L \rangle$ to represent it. Each $p_i \in P$ denotes a PE in the array, and the $\langle p_i, p_j \rangle \in L$ means PE p_j can use the data produced by PE p_i directly through connection network in the CGRA. For a definite II, the CGRA graph is duplicated II times in the time dimension which forms a new directed graph $S_{II} = \langle T_{II}, Q_{II} \rangle$. Modulo scheduling ensures each operation in V being placed into one PE in S_{II} while maintaining feasible data route. Figure 1 shows an example of modulo scheduling on CGRA.
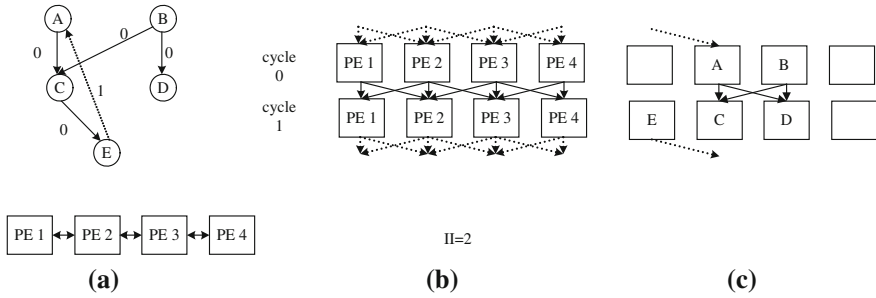
Fig. 1 Example of modulo scheduling on CGRA: **a** loop kernel and target CGRA, **b** the S_{II} with II = 2, and **c** mapping result

The module scheduling on CGRA can be considered as determining the minimal II, with the valid the function $f : V \rightarrow T_{\text{min}II}$ and $g : e \rightarrow 2^{Q_{\text{min}II}}$. The two functions associate operations with PEs meanwhile define route path for each data dependency. We note that the integer linear programming (ILP) can be applied in this problem, but the running time of ILP becomes intolerable if the size of DFG increases. The techniques used in modulo scheduling for VLIW cannot be adopted directly in the mapping problem on CGRA, for they have different constraints and requirements due to architecture divergence. So, a heuristic algorithm for modulo scheduling on CGRA is urgently needed to obtain optimized mapping result of loop kernels in a reasonable time.

3 GA for Loop Mapping on CGRA

3.1 Encoding Scheme and Fitness Calculation

The genetic algorithm has been proved to be an efficient algorithm in single and multi object scheduling problems [5]. Solution of the problem is viewed as a chromosome. The chromosome is a sequence of gene which indicates the variables' value. In modulo scheduling problem with a definite II, the chromosome is defined as:

$$X = \{p_1, p_2, \ldots p_n\}, p_i \neq p_j, n = |V|, p_i \in \{1, 2, \ldots |P| \times II\} \tag{1}$$

Each operation in $v_i \in V$ is assigned with a number p_i which means we place v_i onto the PE p_i in the duplicated graph S_{II}. The variable $p_i = j \times II + k$ where j is the cycle which p_i belongs to, $k \in \{1, 2, \ldots |P|\}$ is the PE number in the CGRA architecture. The example in Fig. 1c has a chromosome denoted by. This method is a sequential encoding scheme of solution space.

In order to evaluate the fitness of a given chromosome, we calculate the number of routable data dependency in the mapping scheme defined by the chromosome. All the data dependency in DFG should be ensured, so a chromosome fits better if more data dependency can be routed in the S_{II}. However, find the best path for all data dependency in S_{II} is also a time consuming and complex problem. To reduce searching space, we consider data dependency one by one. A priority is associate to each $e = \langle u, v \rangle \in E$ according to the as soon as possible (ASAP) and as later as possible (ALAP) scheduling result of u and v [6]. The value $ASAP(u)$ and $ALAP(u)$ are the earliest and latest time that can start execute. We define the priority of data dependency by:

$$priority(\langle u, v \rangle) = ASAP(v) - ALAP(u) \qquad (2)$$

Equation (2) concerns the freedom of an edge in DFG, the less time span an edge can be fit into, the higher priority it will get. We adopt the maze route technique [7] to find route for single data dependency in priority order until route failure. Then the $fitness(X)$ equals number of correctly routed path. Figure 2 illustrates how the fitness is calculated.

Suppose there are two edges in DFG and the operations A, B, C, D's location is represented by $X = \{1, 3, 11, 9\}$. We first route data dependency A to C because of its higher priority. The route makes a PE from 'idle' to 'occupied' as showed in Fig. 2c. Then, the route for B–D fails due to resource confliction. Only 1 path is successful established, so the $fitness(X) = 1$.

3.2 GA Based Module Scheduling Methodology

The algorithm we proposed is depicted in Algorithm 1 as below. We first perform ASAP and ALAP scheduling of the loop kernel to get the priority of each data dependency (line 1–3). We also find the minimal II requirements [8] due to resource constraint and data dependency between different iterations (line 4). Then
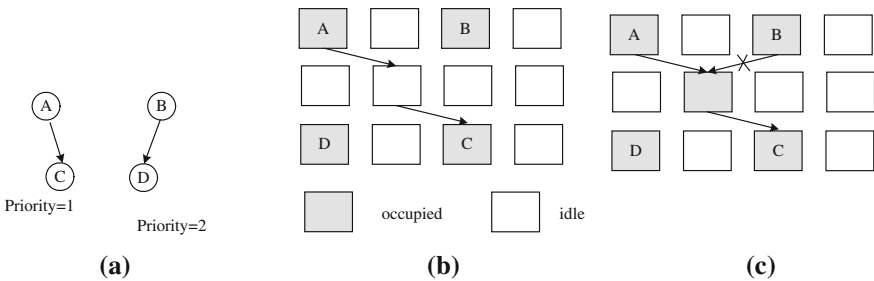
Fig. 2 Fitness calculation by data dependency routing: **a** data dependency, **b** maze route A–C, and **c** maze route B–D
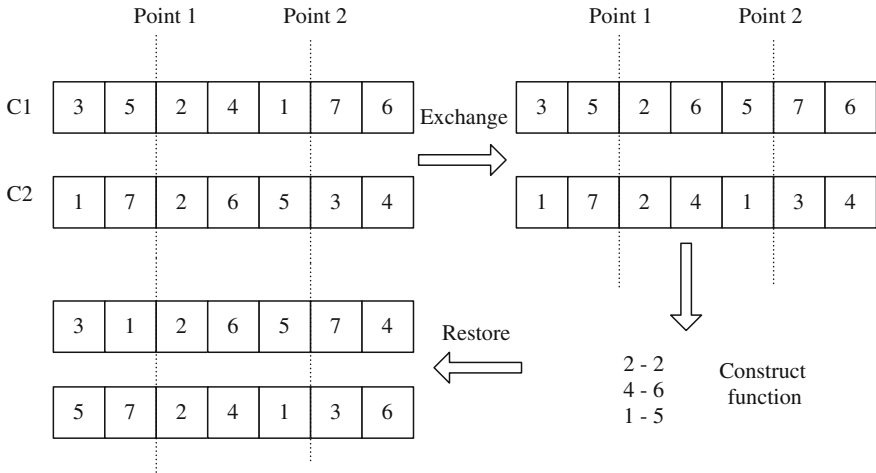
Fig. 3 Crossover operation on chromosomes

the module scheduling start from minII to find valid solution. If it is not obtained after several generations of the genetic evolution, we increase II by 1 and continue to find solution under the new II (line 6–20).

The genetic evolution is started with random population (line 7). It calculates the fitness of individual chromosome one after another by counting route able path in maze route (line 9–15). After that, a number of chromosomes are selected by roulette approach. The higher fitness a chromosome gets, the higher probability it will be selected. Then crossover and mutations are done on the selected chromosomes to form a new population.

The mutation is done through swap the gene in 2 randomly selected positions. It is worth noting that the crossover of two chromosomes may generate an invalid solution. We apply a partially mapped crossover method in this stage to ensure valid of chromosomes. This procedure includes 4 steps:

(1) Choose two points in the chromosomes which divide each into 3 parts;
(2) Exchange the middle part;
(3) Construct function according to exchanged part;
(4) Restore the rest parts invalid chromosome according to the function.

Figure 3 illustrates the procedure. If restoration is not made after exchange, there will be redundant gene in C1 and C2. By replacing duplicated gene through the function, we obtain 2 new valid chromosomes and continue evolution.

Algorithm 1: GA based modulo scheduling for CGRA mapping

1:   perform ASAP and ALAP to the loop kernel;
2:   **for** each $e = \langle u, v \rangle \in E$
3:     calculate the priority using (2);
4:   **end for**
5   $II = minII$;
6:   **while** valid mapping not found
7:     generate random population $M = \{X_1, X_2, ... X_M\}$;
8:     **while** maximum generation is not reached;
9:      **for** each X_i in M
10:        **do**
11:         start maze route for $e = \langle u, v \rangle \in E$ whose priority is the highest;
12:         remove e from E;
13:        **while** maze route success;
14:        $fitness(X_i) = $ number of successfully routed path;
15:      **end for**
16:      select a sub-population by roulette according to the fitness;
17:      generate a new population M by crossover and mutation of the chromosome;
18:     **end while**;
19:    $II = II + 1$;
20:   **end while**

---

4 Experiments and Observation

4.1 Experimental Setup

We evaluate the proposed algorithm using typical target CGRA architecture. The architecture is composed of 4×4 reconfigurable PEs with mesh-like connection networks. The DFGs to be mapped are extracted for 3 applications: WCDMA channel decoder, MIMO-OFDM decoder, LTE channel estimation. A total of 207 DFGs are extracted from these applications with the number of operations per DAG ranging from 25 to 112. We compare the proposed GA module scheduling algorithm with other 2 methods: the integer linear programming (ILP), simulated annealing (SA) algorithm [3]. All experiments were run on an Intel Xeon 6-core CPU with a clock frequency of 2.8 GHz and 8 GB of memory

4.2 Comparison of Mapping Quality

Table 1 shows the best solutions obtained for each mapping algorithms. The performance obtained by ILP is considered as reference point which is optimal. It demonstrates that GA get better result than SA and it is near optimal. A smaller II

Table 1 Comparison of the mapping quality measured in the total II count

Loop kernel	Total II		
	ILP	SA	GA
WCDMA channel decoder	53	69	61
MIMO-OFDM decoder	68	85	74
LTE channel estimation	95	131	102

Table 2 Comparison of the running time

Loop kernel	Running time (s)		
	ILP	SA	GA
WCDMA channel decoder	247,134	57,951	48,425
MIMO-OFDM decoder	395,901	70,237	59,548
LTE channel estimation	427,129	87,395	77,301

mean execution loops more efficiently to accelerate applications. In fact, GA is able to identify and recombine good substructures for further optimization, partial sequence of chromosome is often identified that can lead to a good solution. Thus, the result can be improved in our algorithm.

4.3 Comparison of Mapping Quality

Table 2 shows the running time obtained for each mapping algorithms on the CPU mentioned above. We can see that ILP is time consuming for its full search space. The SA and GA are much faster in a magnitude in finding solutions but they are also long running algorithms. However, the proposed GA takes less time than SA due to its effective route policy for data dependency which ensures high priority edges first. This heuristic reduces the time of routing space search in module scheduling.

5 Conclusion

In this paper, we presented an genetic based modulo scheduling algorithm for mapping loop kernels onto CGRA. Given a kernel's DFG and the target CGRA architecture, the proposed algorithm can find an optimized mapping result with routed data in a reasonable time. The experimental results show that our algorithm outperforms other heuristic methods both in solution's quality and mapping time.

References

1. Zain ul A, Svensson B (2009) Evolution in architectures and programming methodologies of coarse-grained reconfigurable computing. Microprocess Microsyst 33:161–178
2. Shields CO Jr (2001) Area efficient layouts of binary trees in grids. Ph.D. thesis, The University of Texas at Dallas
3. Mei B, Vernalde S, Verkest D, De Man H, Lauwereins R (2002) DRESC: a retargetable compiler for coarse-grained reconfigurable architectures. In: Proceedings of international conference on field-programmable technology, FPT, pp 166–173
4. Park H, Fan K, Mahlke SA, Oh T, Kim H (2008) Edge-centric modulo scheduling for coarse-grained reconfigurable architectures. In: Proceedings of the international conference on parallel architectures and compilation techniques, PACT, pp 166–176
5. Back T, Hammel U, Schwefel H-P (1997) Evolutionary computation: comments on the history and current state. IEEE Trans Evol Comput 1(1):3–17
6. Hwang C-T, Lee J-H, Hsu Y-C (1991) A formal approach to the scheduling problem in high level synthesis. IEEE Trans Comput Aided Des Integr Circuits Syst 10(14):464–475
7. Elghazali M, Areibi S, Grewal G, Erb A, Spenceley J (2009) A comparison of hardware acceleration methods for VLSI maze routing. In: Proceedings of IEEE toronto international conference on science and technology for humanity (TIC-STH), pp 563–568
8. Ramakrishna Rau B (1994) Iterative module scheduling: an algorithm for software pipelining loops. In: Proceedings of the 27th annual international symposium on microarchitecture, pp 63–74

The Investigation and Realization of IPSec Strategy Based on Linux for IPv6

Rui Su and Wei Su

Abstract IPSec is IP security protocol made by IETF which is designed for the next generation network, and is the mandatory part of IPv6 protocol stack. In this paper, the IPSec protocol and the Netfilter mechanism of Linux are introduced briefly. Then a new design idea and system design scheme is presented in detail, the final realization of IPSec protocol which is based on the Netfilter mechanism shows this scheme properly, efficiently and stability, which sets up the basis for the development of security route, VPN, next generation network.

Keywords IPSec · IPv6 · Linux · Netfilter

1 Introduction

To enhance Internet Security, Internet IETF developed a set of protocols, called IPSec described in RFC2401 [1]. IPSec is an architectural system to provide security services in IP layer, and consists of a series of RFC standard protocols, [2–9] which provides connectionless integrity, data origin authentication, protection against replays, confidentiality, (limited) traffic flow confidentiality.
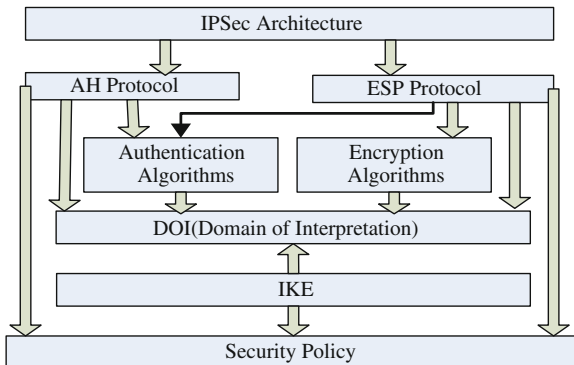
IPSec architectural system is as follows (as shown in Fig. 1). IPSec consists of a couple of separate protocols, as described below:

R. Su (✉) · W. Su
Jining Medical University, Shandong, China
e-mail: suruiyaya@yahoo.com.cn

W. Su
e-mail: suwei9604@sohu.com

Fig. 1 IPSec architectural
system

Authentication Header (AH): described in RFC2402 [2], which provides connectionless integrity and data origin authentication for IP datagrams, and provides protection against replays.

Encapsulating Security Payload (ESP): described in RFC2406 [3], which provides confidentiality, data origin authentication, connectionless integrity, an anti-replay service (a form of partial sequence integrity), and (limited) traffic flow confidentiality.

Internet Key Exchange (IKE): described in RFC2409 [6], which solves the most prominent problem in the setup of secure communication: the authentication of the peers and the exchange of the symmetric keys.

IP Security (IPSec) architecture is described in RFC2401. IPSec consists of an IP packet processing part and a key exchanging part. IPSec processes a packet based on IPSec Security Policy (SP) and IPSec Security Association (SA). SP defines characteristic of security communication between two entities, and uses different protocols under the different conditions, and indicates which packet should be applied with IPSec, bypassed or dropped. SA determines protect what, how to protect and who protect transfer data. SA is the reification and instantiation of SP.

IPSec protocol not only can protect an entire IP payload, but also the top protocol (e.g., TCP/UDP). They are implemented by two different modes of IPSec. The transfer mode is protected the top protocol, and is used for End-to-End communication. The tunnel mode is used to protect entire IP packet, and is mainly used for Security Gateway (SWG) to Security Gateway (e.g., VPN). Under IPSec transfer mode, the special IPSec header is inserted between the original IP header and the top protocol. Under IPSec tunnel mode, the protected original IP packet is encapsulated into another IP packet, and then, an IPSec header is inserted between the outer IP header and the inner IP header (original IP header). AH and ESP protocols can apply under IPSec transfer mode and IPSec tunnel mode at the same time [10–12]. The IP packet structure is as shown in Fig. 2.
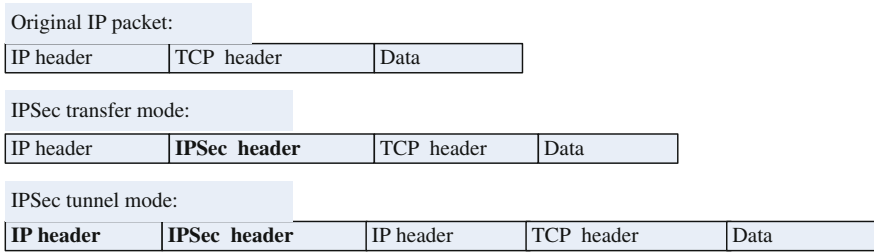
Original IP packet:

IP header	TCP header	Data

IPSec transfer mode:

IP header	**IPSec header**	TCP header	Data

IPSec tunnel mode:

IP header	**IPSec header**	IP header	TCP header	Data

Fig. 2 The IP packet structure

2 Linux Kernel and Netfilter

Linux kernel has been included the IPv6 basic feature set since version 2.2. But IPv6 IPSec was lacking in it. Since Linux kernel 2.5 series, new networking architecture was introduced, called XFRM. In the latest version kernel 2.6 provides stronger support for IPSec than before by using XFRM. Linux kernel 2.6 not only supports AH protocol and ESP protocol, and they depend on authentication algorithms and encryption algorithms are provided by Cryptographic API [13] included in kernel 2.6, but also supports IP Payload Compression Protocol (IPComp [7]) to improve transfers performance.

Linux kernel has been provided a new firewall mechanism, called Netfilter since version 2.4. Netfilter is a new abstract and universal functional framework that allow for packet filtering, network address [and port] translation (NA[P]T) NAT, and other packet mangling. Netfilter is a set of hooks inside the Linux kernel that allows kernel modules to register callback functions with the network stack. A registered callback function is then called back for every packet that traverses the respective hook within the network stack. Netfilter defines five hooks in various points in IPv6 stack. By registering different user-defined IP packet mangling modules on five hooks we can implement IPv6 IPSec. The IPv6 traversal diagram is as shown in Fig. 3.

Fig. 3 The IPv6 traversal diagram
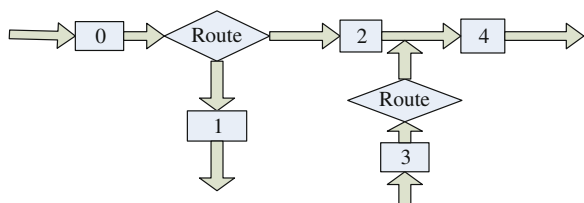
3 System Design

The design thought of overall system makes IP realization and IPSec realize fully integrated through transforming Linux IP protocol stack source code, which uses the Netfilter HOOK mechanism in receiving and sending in the process of IP packet at the appropriate processing position which is called related IPSec processing module [14–16].

According to Netfilter framework mechanism and IPSec architecture on IPv6, we can implement IPSec on IPv6 stack in Linux kernel 2.6. The new packet mangling flow on IP layer is as follows (as shown Fig. 4).

The realization of the system is encapsulation type using in the Linux Netfilter framework HOOK point registered our own IPSec processing module, the IPSec module and IPv6 IP layer seamless fuses in together, so as to realize the IPSec IPv6 environment [17, 18].

When data received by NIC enters into kernel, 'IP layer receiving mangling module' in kernel deals with those data, and gets IP packet. Then it go into 'IPSec Entering Policy' module on hook 0 (NF_IP6_PRE_ROUTING), which concludes that some IP packets can enter local host, and some need to be dropped. Those that will enter into local host are sent into 'Route' module, and then they will be determined to enter into local host or go on to transfer by IP header.

On the hook point 1 (NF_IP6_LOCAL_IN), IP packet that has been transacted by kernel are transferred into 'IPSec entering management module'. This module can distinguish IPSec packet from non-IPSec packet. Non-IPSec packet sent into local host are going to transfer into TCP module directly; And IPSec packet will be
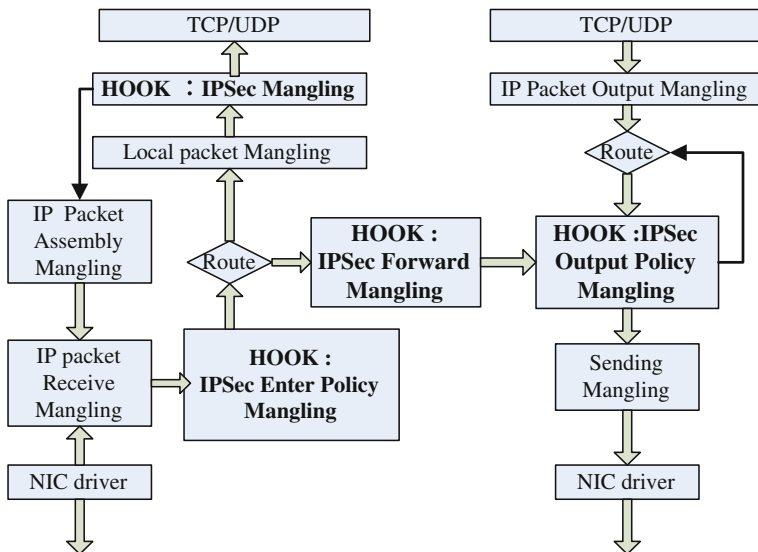

Fig. 4 The new packet mangling flow on IP layer

authenticate or decrypt, get rid of IPSec header, and then it is reassembled in 'IP Packet Assembly Mangling' module. So it will be re-routed to where to go.

For forward packet, IP packet will enter into 'IPSec Output Policy Mangling' module on hook 4 (NF_IP6_POST_ROUTING), which decide to deal with IP packet according to output policy. If IP packet needn't to deal with IPSec, it will transfer into Send Mangling Module. Otherwise, it will be encrypted or authenticated. And then IPSec packet is sent into 'Rout' module to create new route information, because IP packet route information will be modified in the tunnel mode. Finally, IP packet enters into 'Sending' module which sends message into destination host.

Data from transfer layer translates into IP packet in local host. And then route information is inserted into IP packet to determine route export. After routing, IP packet is transferred into 'IPSec Output Policy Mangling' module on hook 4 (NF_IP6_POST_ROUTING). The next progress is the same as the above what said.

The actual development is in the PC machine, then the IPSec module compiler to Linux system kernel, finally transplanted into IPv6 security on a router. The system test first is in PC machine on the gateway tunnel model test. After testing and transplanted to the IPv6 router, and then to test,then the test is successfully passed. This is the system of unique and innovative [19–22].

4 System Module

The system of IPSec implementation mainly consists of five modules as follows (as shown in Fig. 5).

Security Configuring Module: configures SP mostly and saves into SPD; configures SA by hand and saves into SAD.IKE Daemon Program: dynamically creates SA and saves into SAD. IPSec Input Module: handles IP packet that will be
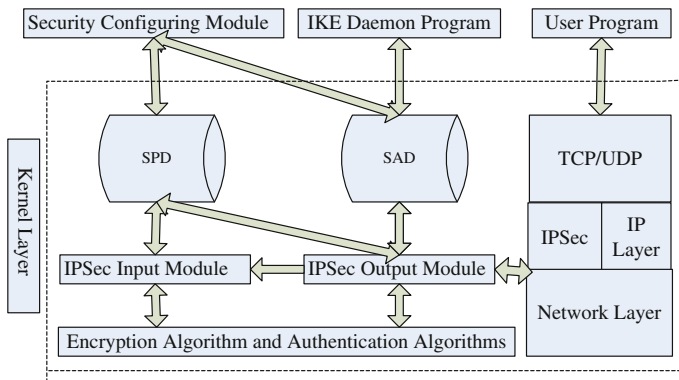

Fig. 5 Five modules of IPSec implementation

entered into IP layer. IPSec Output Module: handles IP packet that will be entered into IP layer. Algorithm Library: provides encryption algorithms and authentication algorithms for IPSec handling.

5 Conclusion

The IPSec implementation in IPv6 based on Netfilter mechanism improves the handling efficiency of IPSec protocol and transfer speed on the IPSec gateway. The simple, efficacious and extendable way makes the protocol of IPSec come true, and ensures the safety of the internet communication. At the same time, the application of IPSec is very wide (e.g., Security Route, Security Gateway, VPN).

With IPv6 technology promotion and IPv6 network construction, which is greatly promotes the development of IPSec. The IPSec application has a grander prospect. Therefore, under the environment of IPSec, the IPv6 network security access technology in the world will get a promotion and widespread application.

References

1. Security architecture for the internet protoco RFC2401
2. IP authentication header RFC2402
3. IP encapsulating security payload RFC2406
4. The internet IP security domain of interpretation for ISAKMP RFC2407
5. Internet security association and key management protocol RFC2408
6. The Internet Key Exchange(IKE) RFC2409
7. IP security document roadmap RFC2411
8. The OAKLEY key determination protocol RFC2412
9. PF_KEY key management API, Version 2 RFC2367
10. The use of HMAC-MD5-96 within ESP and AH RFC2403
11. The use of HMAC-SHA-1-96 within ESP and AH RFC2404
12. The ESP DES-CBC Cipher algorithm with explicit IV RFC2405
13. The NULL Encryption algorithm and its use with IPSec RFC2410
14. Doraswamy N, Harkins D (2000) IPSec: a new generation of network safety standards, Beijing
15. Gao X, Tang J (2004) The realization of the Linux IPSec protocol using Embedded mode, Computer application vol 24
16. FreeS/WAN. http://www.freeswan.org/LinuxFreeS/WAN Project
17. http://packetstorm.widexs.nl/unix/firewall/ipchains/netfilter/netfilter-howto.html
18. Satchell ST, Clifford HBJ (2000) Linux IP Protocol stack source code analysis, Beijing
19. Sklower K (1991) A tree-based packet routing table for Berkeley Unix[A]. In: USENIX Winter 1991 Technical Conference[C], pp 93–104
20. Openswan. http://www.Openswan.org/Linux openswan project
21. Xu X (2002) The analysis and implementation of IPSec protocol, Zhejiang university
22. Xujing,Dianfu Ma (2001) The design and implementation of IPSec, J Beijing univ aeronaut astronaut, 27(4)

An Improved Search Strategy for 8-Digits Puzzle

Aili Han, Zhen Li and Feilin Han

Abstract The search strategy for the 8-digits puzzle is discussed. The state space of the 3-digits puzzle consists of one odd-state-loop and one even-state-loop. For any instance of the 8-digits puzzle, we reduce each state-loop of 3-digits sub-problem to one class-node. If there is a common state between two state-loops, an edge is added between the two class-nodes, and the common state is attached to the edge as its weight. Based on this, we designed an improved method to make the movement in state-loop be compressed from different layers to one layer so that the value of $g(n)$ decreased and the proportion of heuristic function $h(n)$ in evaluation function $f(n)$ increased. The running results show that the efficiency of algorithm has been improved.

Keywords Problem solving · Artificial intelligence · Algorithm analysis · 8-digits puzzle · Heuristic search

1 Introduction

The intelligent searches are usually classified as searches in state space or searches in problem space [1]. The 8-digits puzzle is a typical example of searching in state space. The puzzle randomly places numbers 1–8 on a board of 3*3 cells and one

A. Han (✉) · Z. Li
Department of Computer Science, Shandong University(Weihai), Weihai 264209, China
e-mail: hanal@sdu.edu.cn

Z. Li
e-mail: lizhensd1988@gmail.com

F. Han
College of Software, Shandong University, Jinan 250101, China
e-mail: hanfl9253@163.com

cell remains empty, and then move one number in each step so as to arrange the numbers in order, or illustrate they cannot be arranged in order [2]. An arrangement of 8 numbers on a board is called a state of the 8-digits puzzle. All states make up the state space of the puzzle. If each state in the state space is considered as a node, the whole state space is an undirected graph.

The state space of the 8-digits puzzle is complicated, which has 9! States [3]. Solving this kind of problem usually adopt heuristic search, and the average time complexity is very relevant to the search strategies. For any instance of the 8-digits puzzle, directly using heuristic search cannot avoid searching a large mount of state-nodes, so the efficiency is not always satisfactory.

In this paper, we designed an improved search strategy for the 8-digits puzzle. It makes the movement-in-loop be reduced from different layers to one layer so as to decrease the value of $g(n)$ and increase the proportion of $h(n)$ in the evaluation function. The running results show that the efficiency has been improved.

2 State Space of 8-Digits Puzzle

For any instance of the 8-digits puzzle, the blank may locate at any cell of the board. No matter where it locates, a board of 2*2 cells can always be found to let the blank locate at one cell of it. As shown in Fig. 1, the blank can be regarded as locating at any of the four 3-digits instances. In the following, we discuss the search strategy for the 8-digits puzzle through analyzing the state-change-rule of 3-digits sub-problem.

2.1 State Space of 3-Digits Sub-problem

Each state of the n-digits puzzle can be regarded as a state permutation of 0–n (0 represents the blank). The inversion number [2] in a state permutation refers to the number of digit-pairs that the former is larger than the latter in the state permutation excluding 0. Take the instance shown in Fig. 1 as an example, the state permutation is 123405678, and the inversion number is 0 because the number of digit-pairs that the former is larger than the latter in 12345678 is 0. Obviously, each state corresponds to one state permutation, and vice versa.

The number of all permutations of 0–n is $(n + 1)!$, which is the number of the states in the state space of the n-digits puzzle, so there are $4! = 24$ states in

Fig. 1 An instance of the 8-digits puzzle

1	2	3
4		5
6	7	8

the state space of the 3-digits sub-problem. For all permutations of n numbers, the variable number system can be used to implement no-collision *hash* [4, 5] of each permutation, so we use the variable number system to implement no-collision *hash* of each state of the 3-digits sub-problem.

For the 3-digits sub-problem, when moving from one state to another, the parity of the sum of the inversion number in the state permutation and the line number of the blank keeps unchanged. All the states of the sum being odd form one connected branch; all the states of the sum being even form another connected branch. The number of nodes in each connected branch is $4!/2 = 12$, respectively. Each connected branch is a state-loop that consists of 12 states since the blank in the 3-digits sub-problem can only move along two directions.

Definition 1 In the state space of the 3-digits puzzle, the connected branch in which the sum of the inversion number and the line number of the blank being odd is called the odd-state-loop; the connected branch in which the sum being even is called the even-state-loop.

The odd-state-loop of the 3-digits puzzle is shown in Fig. 2, and the even-state-loop is shown in Fig. 3. For the odd or even state-loop of the 3-digits puzzle, the upper limit of the least number of steps from one state to another is 6. That is, starting with any original state, through at most 6 steps we can arrive at any destination state in the same state-loop.

Definition 2 For the 3-digits puzzle, given an original and a destination states, the state in the preceding four states of the odd-state-loop whose blank has the same location as the original state is called the corresponding state of the original state.
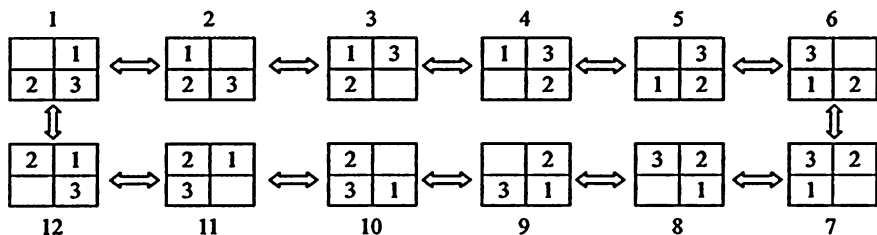

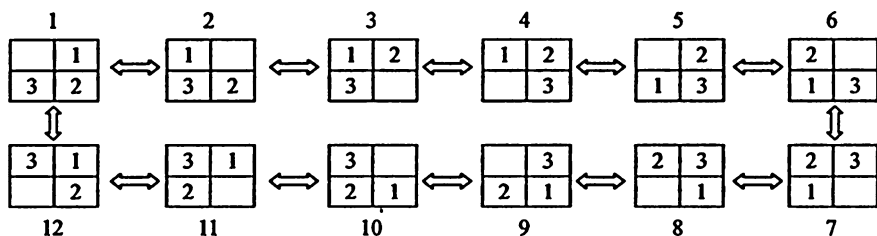Fig. 2 The odd-state-loop of the 3-digits puzzle

Fig. 3 The even-state-loop of the 3-digits puzzle

Fig. 4 An original state and a destination state

2	1		2	
	3		3	1

Fig. 5 The corresponding states of the original and destination states shown in Fig. 4

1	3		1	
	2		2	3

The state obtained from the destination state based on the corresponding relation between the original state and its corresponding state is called the corresponding state of the destination state.

Take the states shown in Fig. 4 as an example. The original and destination states are 2103 and 2031. In the preceding four states of the odd-state-loop of 3-digits puzzle, the state whose blank has the same location as the original state is 1302, that is, the corresponding state of the original state is 1302. The corresponding relation between the original state and its corresponding state is: 2-1, 1-3, 3-2. Based on this corresponding relation, we know that the corresponding state of the destination state is 1023, as shown in Fig. 5.

Take the original state 2310 and the destination state 3021 as another example. In the preceding four states of the odd-state-loop, the state whose blank has the same location as the original state is 1320, that is, the corresponding state of the original state is 1320. The corresponding relation is: 2-1, 3-3, 1-2. Based on this relation, the corresponding state of the destination state is 3012. Note that if the original and destination states locate at different state-loops, there is not need to give its corresponding state since the instance has not solution.

Theory 1 For the 3-digits puzzle, if the original and destination states locate at the same state-loop, the moving process from the original state to the destination one are the same as that from the corresponding state of the original state to the corresponding state of the destination state.

Proof For any instance of the 3-digits puzzle, if the original and destination states locate at the same state-loop, given the same change of the movement on them, when the original state is changed to the destination state, the corresponding state of the original state is changed to the corresponding state of the destination state. That is, the moving steps from the original state to the destination state are the same as that from the corresponding state of the original state to the corresponding state of the destination state, as shown in Fig. 6. □

Based on the above analysis, if the optimal paths starting with one of the preceding four states in the odd-state-loop of the 3-digits puzzle are given, the optimal path starting with any other state can be obtained. Therefore, for any instance of the 3-digits puzzle, the solution can be obtained in the polynomial time through the solutions of the preceding four states.

Fig. 6 The moving steps
from the original state to the
destination state and from the
corresponding state of the
original state to the
corresponding state of the
destination state
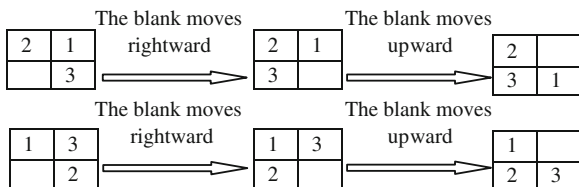
2.2 Relation Between State Spaces of 8-Digits Puzzle and 3-Digits Sub-problem

For any instance of the 8-digits puzzle, when searching for the optimal path in the state space, a large number of computation needs to be done based on different 3-digits sub-problems. For the repeated 3-digits sub-problems, by means of dynamic programming, we compute each sub-problem only once and store the computing result. Thus, the repeated computation can be avoided. For the 8-digits puzzle, we can also obtain two state-loops which are similar with that of the 3-digits sub-problem if we consider the movement of the blank as the movement in the corresponding 3-digits sub-problem.

Definition 3 In the state space of the 8-digits puzzle, the state-loop that has similar structure with that of the 3-digits sub-problem is called the class-state-loop of the 8-digits puzzle.

   Observation 1. The class-state-loop of the 8-digits puzzle has the same characters as the state-loop of the 3-digits sub-problem. Any state of the 8-digits puzzle always locates at one or more class-state-loops.
   Based on the above analysis, we conclude that the state space of the 8-digits puzzle consists of several class-state-loops. To avoid the repeated computation of class-state-loops, we re-organized the state space of the 8-digits puzzle.

3 An Improved Search Strategy for 8-Digits Puzzle

3.1 Basic Operations of State Change

When there is the same state between the class-state-loops in the state space of the 8-digits puzzle, the operations can jump between two loops. Therefore, the operations from one state to another can be classified as follows:

1. Movement in loop: the movement from the original state located at one class-state-loop to the destination state located at the same class-state-loop.
2. Movement between loops: the movement from the original state located at one class-state-loop to the destination state located at another class-state-loop. That

is, jump from one state located at one class-state-loop to another located at another class-state-loop.

3.2 The Improved Search Strategy

Based on the above analysis, we designed an improved organization method of state space for the 8-digits puzzle as follows:

1. Each class-state-loop is compressed into one class-node. The class-node stores all the states in the class-state-loop.
2. If there is one common state between two class-state-loops, an edge is added between the two class-nodes. The weight on the added edge records the common state.

In the improved organization method of state space for the 8-digits puzzle, each class-node includes 12 states, and the common state is attached to the edge between two class-nodes as its weight. When the original and destination states locate at the same class-node, the shortest path can be directly obtained; when the original and destination states locate at different class-nodes, the shortest path can be calculated through the common state on the edge.

Theory 2 In the state space of the 8-digits puzzle, if the original and destination states locate at the same class-state-loop, the shortest path must locate at the class-state-loop.

Proof Suppose that the original and destination states locate at a class-state-loop A, and the shortest path is $s_1s_2\ldots s_{i-1}s_is_{i+1}\ s_{i+2}\ldots s_n$. If the shortest path include one state s_i which do not locate at the class-state-loop A, then according to the organization method of state space, the states s_{i-1} and s_{i+1} must be the same state on the class-state-loop A, that is, there exists one shorter path $s_1s_2\ldots s_{i-1}s_{i+2}\ldots s_n$. This conflicts with that $s_1s_2\ldots s_{i-1}s_is_{i+1}s_{i+2}\ldots s_n$ is the shortest path. If the shortest path include several states $s_is_{i+1}\ldots s_r$ which do not locate at the class-state-loop A, then according to the organization method of state space, the states s_{i-1} and s_{r+1} must be the same state on the class-state-loop A, that is, there exists one shorter path $s_1s_2\ldots s_{i-1}s_{r+2}\ldots s_n$. This conflicts with that $s_1s_2\ldots s_{i-1}s_is_{i+1}\ldots s_rs_{r+1}\ldots s_n$ is the shortest path. Therefore, the shortest path must exclude any state that does not locate at the class-state-loop A.   □

Given any instance of the 8-digits puzzle, if the original and destination states locate at the same class-node, we just search in one class-state-loop. The shortest path from the original state to the destination one is computed by means of the corresponding property of class-state-loop. If the original and destination states locate at different class-state-loops, we need to do the movement between loops. The jump between class-state-loops must pass through the common state on the

edge, which is regarded as that it first jumps to the common state through the movement in loop and then moves in the new loop starting from the common state.

## 4 Heuristic Power of $h(n)$ in the Improved Method

When searching for the shortest path in the state space of the 8-digits puzzle, the key is to determine the next state. If we can make full use of the characteristic information of the 8-digits puzzle to estimate the importance of state when determining the next state, the search process will be more efficient. In the evaluation function $f(n) = g(n) + h(n)$, $g(n)$ denotes the search depth from the original node to node n, and $h(n)$ denotes the estimate cost from node n to the destination node. The heuristic power of $h(n)$ is proportional to the ratio of $h(n)$ to $f(n)$, so increasing the value of $h(n)$ or decreasing the value of $g(n)$ can increase the heuristic power of $h(n)$.

The previous search strategy [1] for the 8-digits puzzle is shown in Fig. 7, where the depth of the states 2 [3] and 1 [3] are 3, i.e. $g(n) = 3$. For the same instance, the improved search strategy presented in this paper is shown in Fig. 8. The improved method makes the movement in loop be compressed from different layers to one layer, which decreases the value of $g(n)$ and thus increases the

Fig. 7 The previous search strategy for the 8-digits puzzle
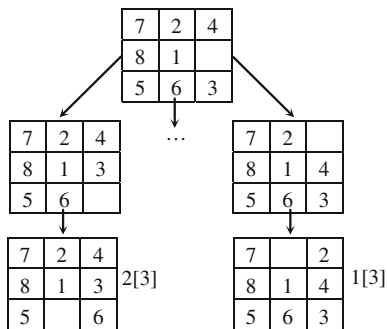
Fig. 8 The improved search strategy for the same instance as Fig. 7
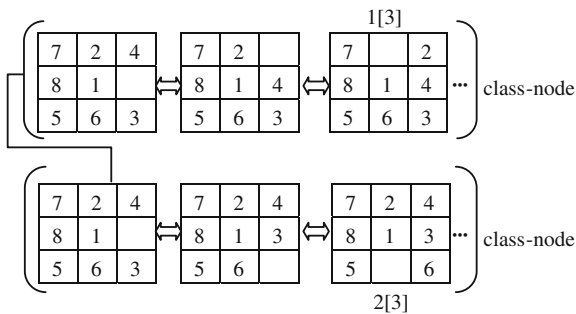
proportion of $h(n)$ in the evaluation function. Compared with Fig. 7, the state 1[3] in Fig. 8 is in the first layer instead of the third layer. When moving from the original state to the state 1 [3], only need the movement in loop, $g(n) = 1$. The state 2 [3] in Fig. 8 is in the second layer instead of the third layer, $g(n) = 2$. In the evaluation function $f(n) = g(n) + h(n)$, the decrease of $g(n)$ results in the increases of the ratio of $h(n)$ to $f(n)$. Thus, the heuristic power of $h(n)$ has increased.

Based on the above analysis, we know that, by means of the improved method, the value of $g(n)$ decreases in the evaluation function $f(n) = g(n) + h(n)$, which results in the increasing ratio of $h(n)$ to $f(n)$. Therefore, the efficiency of solving the 8-digits puzzle has been improved.

5 Results and Discussions

It is well known that, the more the nodes being expanded in the running process of an algorithm, the higher the time complexity of the algorithm. The original heuristic algorithm [1] for the 8-digits puzzle enlarges the value of $h(n)$ to increase the heuristic ability. The formula of computing the estimation function is $f(n) = g(n) + h(n)*k$, $k = 3$. The improved method in this paper increases the heuristic ability through decreasing the value of $g(n)$. It compresses the movements-in-loop from different layers to one layer so as to decrease the value of $g(n)$. The comparison between two algorithms is given in Table 1.

For the original algorithm, the total running time is 7342734 ms, the total effective items is 181439, the total number of steps in finding the shortest paths is 4008894, the total times of leaving queues is 187365910, the total times of entering queues is 296338511. For the improved algorithm, the total running time is 5907015 ms, the total effective items is also 181439, the total number of steps in finding the shortest paths is 5507880, the total times of leaving queues is 44677498, the total times of entering queues is 258001665.

Consider the 181439 effective items. The average running time decreases 19.55 %, the average times of leaving queues decreases 76.15 %, the average times of entering queues decrease 12.94 %. In the 181439 effective items, the number of items whose shortest paths are improved is 22,300, the number of items whose times of leaving queues are improved is 154098, and the number of items

Table 1 Comparison on the running times of two algorithms

Compared items	The original algorithm	The improved algorithm	Ratio
The running time	7342734 ms	5907015 ms	80.45 %
The effective items	181439	181439	1
The number of steps	4008894	5507880	137.39 %
The times of leaving queues	187365910	44677498	23.85 %
The times of entering queues	296338511	258001665	87.06 %

whose times of entering queues are improved is 83652. According to the running results of the improved program, the efficiency of our algorithm has been improved, but it still cannot avoid entering the local minimum in the running process of algorithm.

References

1. Lu R (2000) Artificial intelligence. Science Press, Beijing, pp 152–339
2. Han A (2004) A study on the solution of the 9-room problem by the state space method. J Shandong Univ (Sci Ed) 34(4):51–54
3. Huaxian H (1988) Introduction to artificial intelligence. Northwestern Polytechnical University Press, Xi'an, pp 1–10
4. Meng F (2009) Research on learning-based method of solving the 9-room problem and application. National University of Defense Technology, Changsha
5. Cao Z (2005) Breadth-first search of hua rong dao: application of hash search and heuristic search. J Electr Power 20(1):13–23

DNA Sequence Motif Discovery Based on Kd-Trees and Genetic Algorithm

Qiang Zhang, Shouhang Wu, Changjun Zhou and Xuedong Zheng

Abstract In the post-genomics era, recognition of transcription factor binding sites (DNA motifs) to help with understanding the regulation of gene is one of the major challenges. An improved algorithm for motif discovery in DNA sequence based on Kd-Trees and Genetic Algorithm (KTGA) is proposed in this paper. Firstly, we use Kd-Trees to stratify the input DNA sequences, and pick out subsequences with the highest scoring of the hamming distance from each layer which constitute the initial population. Then, genetic algorithm is used to find the true DNA sequence motif. The experiment performing on synthetic data and biological data shows that the algorithm not only can be applied to each sequence containing one motif or multiple motifs, but also improve the performance of genetic algorithm at finding DNA motif.

Keywords DNA sequence · DNA motif · Kd-Trees · Genetic algorithm

1 Introduction

With the launching of the Human Genome Project and the rapid development of modern biotechnology, the growth of bioinformatics data shows the trend of the explosion. Revealing the intrinsic link between huge amounts of data and biological significance is one of the urgent problems now [1].

Genetic Algorithm (GA) is a stochastic search algorithm, which is used to search large, non-linear spaces where professional knowledge is lacking or difficult to encode. In recent years, GA is introduced to solve the problem of motif discovery. Stine et al. [2] proposed the structured genetic algorithm (ST-GA) that can

Q. Zhang (✉) · S. Wu · C. Zhou · X. Zheng
Key Laboratory of Advanced Design and Intelligent Computing, Ministry of Education,
Dalian University, Dalian 116622, China
e-mail: zhangq@dlu.edu.cn

identify the different length of the motifs. Liu et al. [3] introduced an algorithm of Finding Motifs by Genetic Algorithm (FMGA), it used a general GA framework, and designed a genetic operator which can speed up the convergence process. Che et al. [4] solved Motif Discovery problem using a Genetic Algorithm (MDGA), this algorithm can prediction binding sites for homologous effectively, and the fitness is obtained by combination of sites in each column. Congdon et al. [5] developed an algorithm called Genetic Algorithm approach to Motif Inference (GAMI). This method is very effective for the long sequence motif. Lately, Huo et al. [6] presented an original algorithm (GARPS) that combined Genetic Algorithm with Random Projection Strategy which performs better than the projection algorithm. Whereas each type of method has some shortcomings, such as using exhaustive search methods to bring high computation, local search methods converge to local optima value easily, but combination of different methods for motif discovery show good result.

2 The Steps of Algorithm KTGA

In this work, we present an improved algorithm KTGA that optimizes Genetic Algorithm via Kd-Trees to identify DNA motif. The standard genetic algorithm usually starts from an initial population generated at random and attempts to improve on it using genetic operators in the guidance of a fitness function. Randomly generated population will rarely come close to an optimal solution and there is little chance of converging to the optimal solution. Thus, we generate an initial population with a good starting point for the genetic algorithm by using Kd-Trees.

2.1 Kd-Trees

In order to reduce the complexity, we perform a hierarchical clustering preprocessing method based on Kd-Trees. Original Kd-Trees was proposed by Bentley in 1975. This method is an attempt to speed-up the execution of nearest neighbor queries by defining a recursive binary partitioning of a k-dimensional dataset, where the root node contains all data. Most such methods partition the data at each tree level using an appropriate hyper plane perpendicular to the direction. In order to deal with sequential data, Blekas et al. [7] propose a modification on the basis of original Kd-Trees. Starting with a root node that contains all set of subsequences X, at each step they partition the set of subsequences at each node using an appropriate criterion. After calculating the relative frequency values $f_{l,k}$ of each character α_l at each subsequences position k in the subset of subsequences contained in that node, determine the position q that exhibits the largest entropy value H_k:

$$H_k = -\sum_{\substack{\alpha_l \in \sum \\ f_{l,k} > 0}} f_{l,k} \log f_{l,k} \tag{1}$$

$q = \arg \max_{k=1,\dots,w}\{H_k\}$. The partitioning procedure is implemented by initially sorting the characters α_l in the position q according to their relative values $f_{l,q}$ and then marking them as 'left' or 'right'. So two subsets are created, which are successively filled with the substrings that contain the 'left' and the 'right' characters in the position q. An example is shown in Fig. 1 where the second position with the maximum entropy is selected for partitioning.

The above recursive procedure constructs a tree with several nodes and the partitioning for a node is terminated when the number of included subsequences is lower than a fixed value T. Every node of the tree contains a subset of the original set of subsequences. In each node, all of the subsequences scored by hamming distance, then, select the highest score of the subsequence. At last, we get the total set of leaf nodes C which contains the highest score subsequences in each leaf node. Our experiments have shown that this partitioning technique drastically accelerates the global search procedure without affecting its performance.

The above recursive procedure constructs a tree with several nodes and the partitioning for a node is terminated when the number of included subsequences is lower than a fixed value T. Every node of the tree contains a subset of the original set of subsequences. In each node, all of the subsequences scored by hamming distance, then, select the highest score of the subsequence. At last, we get the total set of leaf nodes C which contains the highest score subsequences in each leaf node. Our experiments have shown that this partitioning technique drastically accelerates the global search procedure without affecting its performance.
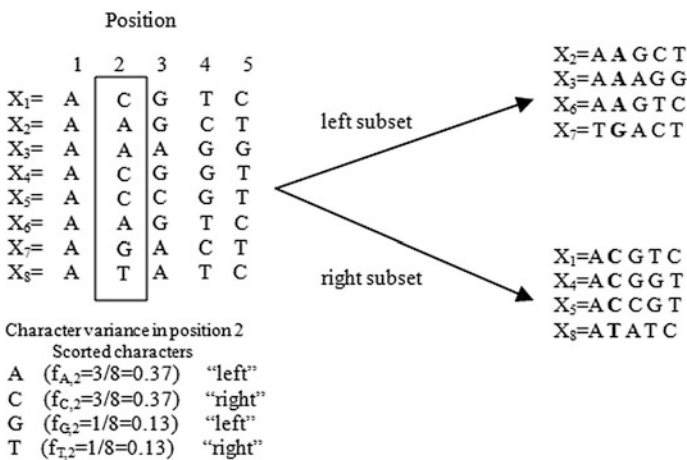

Fig. 1 Partitioning occurs in the second position that presents the maximum character variance

2.2 Adaptive Genetic Algorithm

2.2.1 Individual Representation

The first step in the implementation of genetic algorithm is to generate an initial population. We can use the results from Kd-Trees step. If the number of subsequence produced by Kd-Trees is less than the number of initial population required by genetic algorithm, the algorithm will cycle Kd-Trees step until these two numbers equal. In this paper, we use the real motif expressed sequence as individual coding. For example, the motif sequence 'TCGCAAT' encoded into the sequence of 'TCGCAAT'.

2.2.2 Evaluation Function

In order to assess a consensus sequence, we need to define an evaluation function to compare different consensus, the possible candidate of the motif.

Definition 1: The adaptation value for the motif P_n and the same length of subsequence S_m^i is:

$$FS_m^i\left(S_m^i, P_n\right) = \frac{\sum_{j=1}^{W} match\left(S_m^{i,j}, P_n^j\right)}{W}, \ match\left(S_m^{i,j}, P_n^j\right)$$

$$= \begin{cases} 1 & S_m^{i,j} = P_n^j \ for \ P_n^j \in \{A, T, G, C\} \\ 0 & S_m^{i,j} = P_n^j \ for \ else \end{cases} \tag{2}$$

where: W is the length of the motif P_n, $S_m^{i,j}$ is element on the j position of the subsequence S_m^i, P_n^j is element on the j position of the motif P_n.

Definition 2: In the sequence S_m, we assume that K subsequences of length equal to the motif P_n, such as, $S_m^1, S_m^2, \ldots S_m^k$, then the fitness value of the motif P_n and the sequence S is :

$$FS_m(S_m, P_n) = \sum_{l=1}^{K} FS_m^l\left(S_m^l, P_n\right), \ S_m^l : d\left(P_n, S_m^l\right) \leq t \tag{3}$$

where: t denotes the allowed mismatch number, entered by the user, $d(P_n, S_m^l)$ denotes hamming distance between the motif P_n and the sequence S_m^l.

Definition 3: The fitness value between the motif P_n and the set of sequence $S = \{S_1, S_2, \ldots, S_m\}$ is:

$$TFS(S, P_n) = \frac{\sum_{i=1}^{m} FS_l(S_l, P_n)}{m} \tag{4}$$

cgat	GGATTAC	aggt
ggat	CGATTAC	agat
aggt	GGATTAC	acat
agat	AGATTAC	aggc
agat	TGATTAC	agcg

SHIFT →

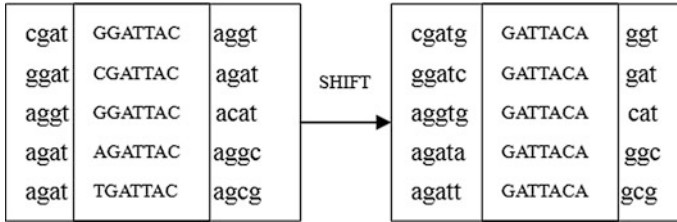| cgatg	GATTACA	ggt
ggatc	GATTACA	gat
aggtg	GATTACA	cat
agata	GATTACA	ggc
agatt	GATTACA	gcg

Fig. 2 Shift operation

2.2.3 Evaluation Function

To create the next generation, we start with the current population and apply the roulette selection to it to create an intermediate population. After evaluating each of the individuals in the current population, selection is carried out. We use Kd-Trees to stratify the current population, and pick the individual with the highest scoring of the hamming distance from each layer until the intermediate population is full. Then we applied crossover and mutation operators to the intermediate population to create the next population. We use the adaptive genetic algorithm with a standard one-point crossover. Crossover is applied to choose individuals with probability between PC_1 and PC_2 randomly and two new individuals are formed that are joined in the next population. After crossover, for each bit of the individuals of the population, we apply a mutation operator with probability PM_1 between PM_2. After the process of selection, crossover and mutation, we use the elitist strategy to choose individuals.

2.2.4 Shift Optimization

The optimal motif identified by genetic algorithm may be displacement of the true motif. In order to eliminate this problem, we add the shift optimization in the KTGA. Shift operation can move all implanted motif instance in two directions (left or right) at the same time. If the fitness value of the motif after moving is greater than before, then replace the motif, the example shown in Fig. 2.

2.2.5 Algorithm KTGA

Algorithm KTGA:
Input N DNA sequences, T, t, PC_1, PC_2, PM_1, PM_2, n
Output   DNA sequence motif
  1: generate initial population
  2: t = allowed mismatches number, n =the number of initial population, T
  3:     do Kd-Trees

4:      if $C \geq n$
5:          end
6:      else
7:          $T = T + 1$
8:      end
9: evaluate individuals in initial population C
10: while stopping criteria is not satisfied do
11: perform roulette selection
12: generate offspring with crossover operator
13: mutate offspring
14: SHIFT operation
15: evaluate offspring
16: perform Elitist strategy
17: end while
18: return the best individual

Lines 1–9 create the initial population of size C from the set of input sequence, obtained from Kd-Trees and evaluate the individuals in the initial population according to the fitness function derived from (6). The size of the population is kept constant during iterations. Lines 10–16 is a while loop, it iteratively first does roulette selection and then does one-point crossover and mutation operators on the intermediate population with probabilities PC_1 to PC_2 and PM_1 to PM_2, respectively, then, use the SHIFT operation to optimize individual. End of the algorithm until the stopping criteria is satisfied.

3 Experimental Results

We test the algorithm on the synthetic data, which contains 10 DNA sequences of fixed length 50 with implanted (15, 4) motif instance in each sequence (Fig. 3). The motif discovery result is as follows. Logo diagram of motif representation in Fig. 4.

Fig. 3 The input sequences

Fig. 4 The motif logo of synthetic data

Table 1 Precision rates of different algorithms on synthetic data with implanted (15, 4) instances [10]

Algorithm	Sequence length				
	100	200	300	400	500
Consensus	0.92	0.90	0.76	0.78	0.64
Gibbs DNA	0.90	0.81	0.51	0.64	0.53
MEME	0.99	0.98	0.98	0.97	0.87
SP-STAR	1.00	0.98	0.97	0.95	0.86
Phylo Gibbs	0.98	0.93	0.90	0.92	0.89
KTGA	1.00	0.99	0.98	0.99	1.00

By the results above, we can see that the motif was found by our algorithm is the same as the known results, and shows that the algorithm is effective.

For the performance evaluation, let K be the set of known signal positions in a sample, and let P be the set of positions predicted by our method. We use the precision rate (performance coefficient)

$$\frac{|K \cap P|}{|K \cup P|} \tag{5}$$

defined by Pevzner and Sze [8] for the evaluation. In Table 1, we compare the precision rate of our method with those of other widely used motif finding algorithms. We derive the average precision rates by executing different algorithms on five simulated data, each of which contains 20 sequences and each sequence of length is between 100 and 500 bps with implanted (15, 4) motif instances.

From the Table 1, we can see the precision rate of our method is over 98 % for sequences of various lengths between 100 and 500, which is better than the precision rate of the compared algorithms. In particular, when the sequence length is larger than 400 bps, the precision rates of the other algorithms decrease quickly except for KTGA. Overall, the present algorithm achieves the best precision rate.

To test practicality of the algorithm, KTGA was tested on datasets of LexA, DHFR and E.coli CRP. LexA database contains 16 DNA sequences with 19 known motif instances. DHFR consist of sequences in the upstream regions of eukaryotic
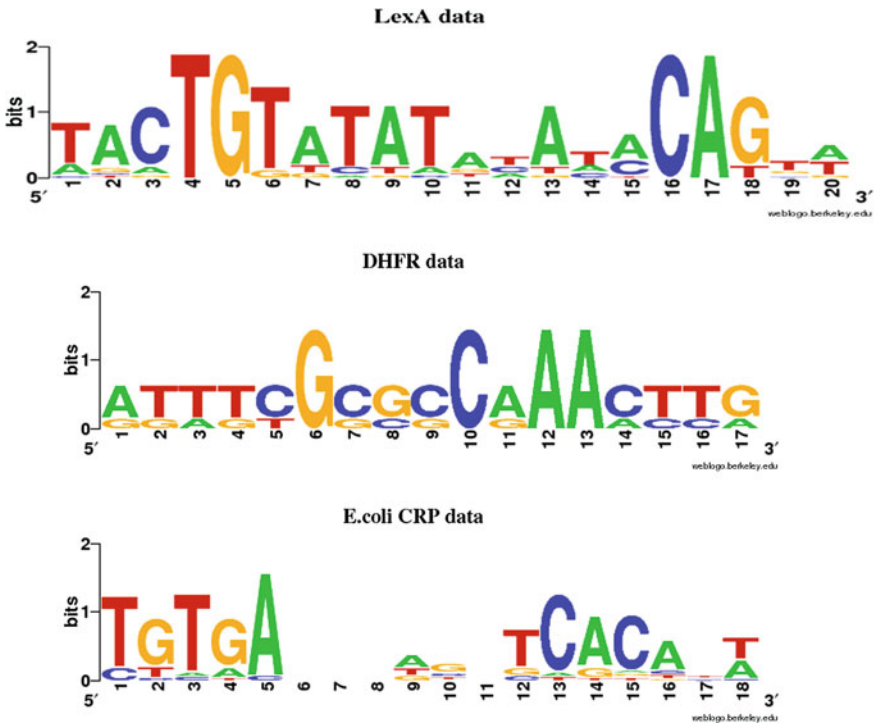
Fig. 5 The motif logo of LexA data, DHFR data and E.coli CRP data

genes. E.coli CRP data form GenBank Release 55, each sequence contains at least one CRP-binding site. By adjusting t and T, discovered motifs are shown in Fig. 5.

In the Table 2, we compared the KTGA algorithm with TreeMotif [9] algorithm on the datasets of LexA, DHFR and E.coli CRP. The correct rates of these two algorithms are almost the same on the datasets of DHFR and E.coli CRP, but on the dataset of LexA, the correct rate of KTGA is 20 % higher than the Tree Motif algorithm. This is because the Tree Motif algorithm assumes that each sequence has exactly one motif occurrence per sequence (OOPS). But, the KTGA algorithm not only can be applied to each sequence containing one motif instance or multiple motif instances.

Table 2 The correct rates of the LexA, DHFR and E.coli CRP datasets

Algorithm	Database	Discovered	Published	Correct rate (%)
Tree motif	LexA	TACTGTGTATATANANNCAGTNNAAT	TACTGTATATATATACAGTA	80
KTGA		TACTGTATATATATACAGTA		100
Tree motif	DHFR	TTCGCGCCAAAACT	ATTTCGCGCCAAACT	86
KTGA		ATTTCGCGCCAAACTTG		88
Tree motif	E.coli CRP	TGTGANNNAGNTCACATT	TGTGANNNGNTCACA	83
KTGA		TGTGANNNAGNTCACATT		83

4 Conclusion

In this paper, a DNA motif discovery algorithm based on Kd-Trees and Genetic Algorithm (KTGA) is proposed. The KTGA algorithm can solve the motif discovery problems in synthetic data and biological data effectively. Kd-Trees is used to stratify the input DNA sequences, and pick the subsequence with the highest scoring of the hamming distance from each layer which constitutes the initial population. Then, the genetic algorithm is used to find the true DNA sequence motif. From the experimental results, we can find that the algorithm not only can be applied to each sequence containing one motif instance or multiple motif instances, but also improve performance at finding DNA motif.

References

1. Ihuegbu NE, Stormo GD, Buhler J (2012) Fast, sensitive of conserved genome wide motifs. J Comput Biol 19:139–147
2. Stine M, Dasgupta D, Mukatira S (2003) Motif discovery in upstream sequences of coordinately expressed genes. Evol Comput 3:1596–1603
3. Liu FFM, Tsai JJP, Chen RM, Shih SH (2004) FMGA: finding motifs by genetic algorithm. In: Proceedings of 4th IEEE symposium on bioinformatics and bioengineering, Taichung, pp 301–309
4. Che DS, Song YL, Rasheed K (2005) MDGA: motif discovery using a genetic algorithm. Genetic and Evolutionary Computation, New York, pp 447–452
5. Congdon CB, Fizer CW, Smith NW (2005) Preliminary results for GAMI: a genetic algorithm approach to motif inference. In: Proceedings of symposium on computational intelligence in bioinformatics and computational biology, CA, pp 1–8
6. Huo HW, Zhao ZH, Stojkovic V, Liu LF (2010) Optimizing genetic algorithm for motif discovery. Math Comput Model 52:2011–2020
7. Blekas K, Fotiadis DI, Likas A (2003) Greedy mixture learning for multiple motif discovery in biological sequences. Bioinformatics 19:607–617
8. Pevzner PA, Sze SH (2000) Combinatorial approaches to finding subtle signals in DNA sequences. Intell Syst Mol Biol 8:269–278
9. Sun HQ, Low MYH, Hsu WJ, Tan CW, Rajapakse JC (2012) Tree-structured algorithm for long weak motif discovery. Bioinformatics 27:2641–2647
10. Huang CW, Lee WS, Hsieh SY (2011) An improved heuristic algorithm for finding motif signals in DNA sequences. IEEE/ACM Trans Comput Biol Bioinf 8:959–975

Footplant Detection Based on Adjusting the Foot Height Threshold Automatically

Qiang Zhang, Jingchao Zhang, Dongsheng Zhou and Xiaopeng Wei

Abstract In this paper, we present a footplant detection method by adaptive determining foot height threshold. In our method, we firstly divide one step into a number of stages, each stage has different types of constraints. Then adaptively electing foot height threshold according to motion type and the length l of the minimum constraint frame that designed by the user and the maximum error φ of l to determine which foot should be planted. Finally, we verify the effectiveness of our approach from the path offset error.

Keywords Footplant detection · Step · Coordinate transformation · Foot height threshold

1 Introduction

When we edit or synthesize the sports, the movement of the feet's end-effecter [1] (tiptoe) plays a very important role on the effect of the motion editor [2] or synthetic, and influence the sport's quality after editing in a certain extent. Sometimes we need to display and specified the constraint position of the tiptoe. For example, the tiptoe of two or several continuous frames is in the same position in space at the original motion sequence. But it will no longer in the same position after motion editing. It appears footskate phenomenon. This makes the movement anamorphic after editing, and destroys the verisimilitude of the original movement.

Foot support means that the tiptoe fixed on the ground in a certain position at a period of time. Keeping foot support could prevent the footskate phenomenon when we edit or synthesize the sports. In order to clear footskate [3–6], we should

Q. Zhang (✉) · J. Zhang · D. Zhou · X. Wei
Key Laboratory of Advanced Design and Intelligent Computing, Dalian University,
Ministry of Education, Dalian, China
e-mail: zhangq@dlu.edu.cn

835

do the constraint detection on the foot support's frames firstly. And then we could make good use of the tiptoe's translation to remove slide.

2 Related Work

Real human motion usually contains variety of constraints [7, 8], for example footplants, which are periods of time when a foot or part thereof remains in a fixed position on the ground. Many works have been dedicated to the footplant.

Liu and Popovic et al. [9] proposed a universal constraint detection methods by adaptively finding a subset of the character's body parts should be fixed in space for some period of time. The number of a positional constraint was considered to be at least n frames. In [10], Glardon et al. introduced a system capable of detecting constraint for a real-time. At first, the system generating future motion postures in preprocessing phase and then selecting height threshold values according to motion type and quality. However, those methods are not reliable for noisy or imperfect data. In order to solve those problems, Gao et al. [11] introduce hardware for footplant detection, he equips the actor with a simple pressure trigger circuit under his shoes. And Ikemoto et al. [12] proposed a method use machine learning to train an oracle, but it wastes a lot of time in training oracle.

In order to combining the advantages of the above method, this paper puts forward a algorithm which could automatically adjust the threshold of the feet height and decide the support leg. According to three aspects: movement type of the capture data, the length l of the minimum constraint frame that designed by the user and the maximum error η of l, we could determine the height threshold ϕ of the feet. If the feet's height of a frame is less than ϕ, then we call this foot as "support leg". Generally, support leg can be divided into 4 types: left foot support, right foot support, double-support, feet without support. We set a movement step as a cycle of motion in the periodic motion. From the first step of data capture, it would adjust ϕ on the basis of η and divide support leg dynamically. There are two main reasons for why we choose dynamic division support leg method. On one hand, the data get from it's interaction with the user meets the user's needs very well. On the other hand, this algorithm always has better result for noise data processing.

3 Preprocessing

3.1 Step

The intuition tells us that for a given type of locomotion (walk or run),a foot joint constraint occurs roughly in side a fixed time interval, regardless of other motion characteristics. So a normal walking can be divided into the following phases (see Fig. 1):
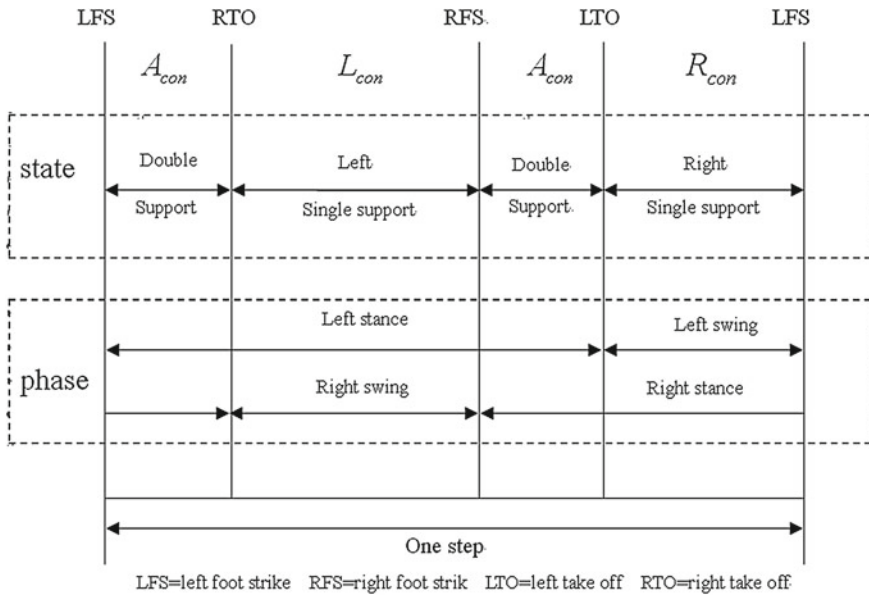
Fig. 1 One step of a walking motion [6]

$$A_{con} - R_{con} - A_{con} - L_{con} - A_{con} \tag{1}$$

And a normal running also can be divided into:

$$R_{con} - N_{con} - L_{con} - N_{con} - R_{con} \tag{2}$$

where A_{con} represents double-support; R_{con} represents right foot support; L_{con} represents left foot support; supported; N_{con} represents feet without support.

3.2 Coordinate Transformation

Because BVH data's structure is hierarchical structure, each node's position actually is relative to its father node's position. If we want to find out each node's location in the global coordinate system, we must find out the transformation relationship between two coordinate system first.

If we fetch a point $p(x_0, y_0, z_0)$ and three Euler Rotation Angle $A(A_x, A_y, A_z)$ in the coordinate system C_0, and then we define a new coordinate system C_1 by the point p and the Rotation Angle sequence. We assume that the Euler Angle's rotation order is z, x, y, and then the transformation matrix from the coordinate system C_0 to the coordinate system C_1 is as below:

$$M = T * R \tag{3}$$

Among them, T, R respectively means that:

$$T = \begin{pmatrix} 1 & 0 & 0 & o_x \\ 0 & 1 & 0 & o_y \\ 0 & 0 & 1 & o_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad (4)$$

$$R = R_z(A_z) * R_x(A_x) * R_y(A_y) \qquad (5)$$

And among the above formula:

$$R_z(A_z) = \begin{pmatrix} \cos A_z & -\sin A_z & 0 & 0 \\ \sin A_z & \cos A_z & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} R_y(A_y)$$

$$= \begin{pmatrix} \cos A_y & 0 & \sin A_y & 0 \\ 0 & 1 & 0 & 0 \\ -\sin A_y & 0 & \cos A_y & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} R_x(A_x) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos A_x & -\sin A_x & 0 \\ 0 & \sin A_x & \cos A_x & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$(6)$$

Here we take BVH data as an example to find out the left tiptoe's coordinate in the global coordinate system. The hierarchical structure from the BVH root node to the

left tiptoe is that:

```
HIERARCHY
ROOT hip
{
  OFFSET 0.00000 0.00000 0.00000
  CHANNELS  6  Xposition  Yposition  Zposition  Zrotation
Yrotation Xrotation
        JOINT lhipjoint
        {
          OFFSET 0 0 0
          CHANNELS 3 Zrotation Yrotation Xrotation
          JOINT lfemur
          {
            JOINT ltibia
            {   ...
              JOINT lfoot
              {   ...
                JOINT ltoes
                {
                  OFFSET 0.17032 -0.46794 2.08086
                  CHANNELS  3  Zrotation  Yrotation
Xrotation
}
```

lfoot is the father node of *ltoes*. After translating x, y, z according to the corresponding *OFFSET* and rotating (*Xrotation, Yrotation, Zrotation*) based on the coordinate system *lfoot*, we get coordinate system *ltoes*.

The transformation matrix between them is M_{ltoes}

$$M_{ltoes} = T_{ltoes}(OFFSET) * R_z(Zrotation) * R_y(Yrotation) * R_x(Xrotation) \qquad (7)$$

So if we fetch an any point P whose coordinate is V_{lfoot} from the coordinate system *lfoot*, and its corresponding point coordinate is V_{lfoot} in the son coordinate system *ltoes*.

$$V_{lfoot} = M_{ltoes} * V_{ltoes} \tag{8}$$

Recurring upgrade gradually, we can get the coordinate of this point in the global coordinate system V_{world}:

$$V_{world} = M_{Hip} * M_{lhipjoint} * M_{lfemur} * M_{ltibia} * M_{lfoot} * M_{ltoes} * V_{ltoes} \tag{9}$$

4 Adjust the Threshold of the Foot Height Adaptively

The whole train of thought of this algorithm which aims at deciding support leg through adjusting the threshold of the foot height adaptively is that: using movement type of the capture data, the length l of the minimum constraint frame that designed by the user and the maximum error η of l those three aspects to determine the height threshold ϕ of each foot step. If the foot's height of a frame is less than ϕ, then we call this foot as "support leg".

The tiptoe of our normal person may swing in the athletic process. It leads to the incorrectness on the support leg by using the height of the tiptoe (see Fig. 2). Here we will use the height of the ankle instead of the height of the tiptoe to get our result.

Based on the method that using unified feet height to determine support leg which was raised by Liu [9] etc., we put forward the adjusting foot height threshold adaptively method. Steps in detail are as follows:

(1) Preprocessing: calculate the coordinates of the left and right ankle in the global coordinate system. That is corresponding V_{world} of V_{lfoot} and V_{rfoot}.
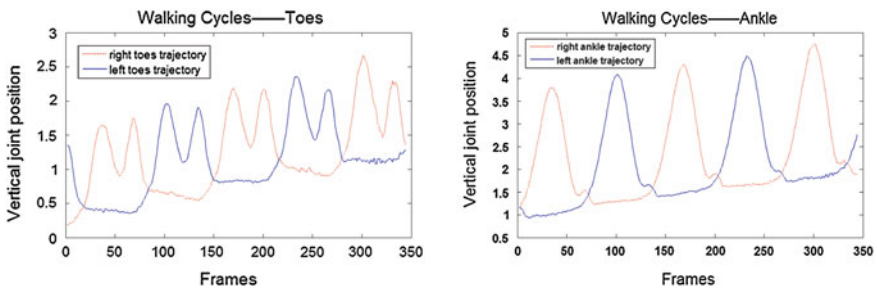
Fig. 2 Compare toes with ankle. *Top* the *vertical trajectory* of the toes; *Bottom* the *vertical trajectory* of the ankle
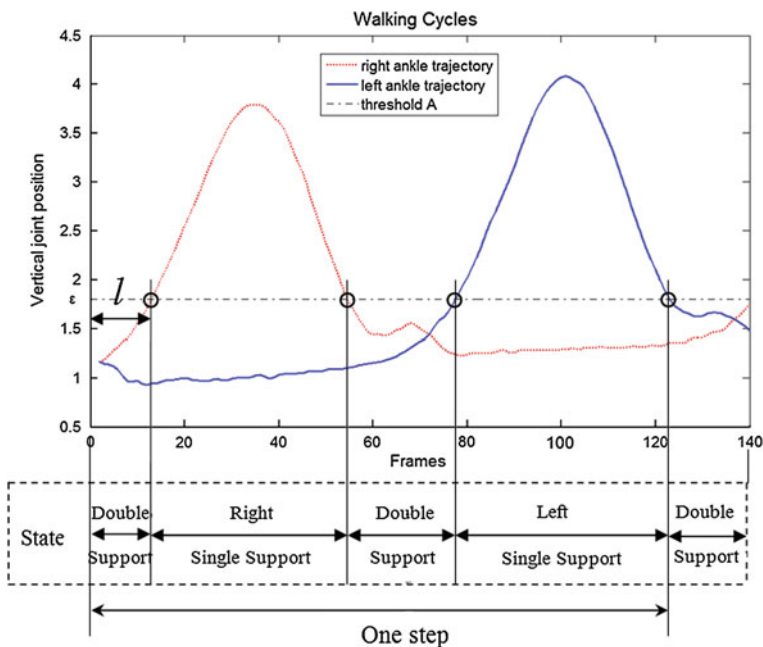
(2) Find out the ankle height threshold value of the first step by using the movement type and the length of minimum support leg constraint frame set by the user.

Different movement type can satisfy different support leg constraint. The support leg walks in one step meets $A_{con} - R_{con} - A_{con} - L_{con}$. Combined with the length l of minimum support leg constraint frame set by the user, we could find out the constraint height ε of the first ankle step (see Fig. 3).

(3) According to the height threshold ε of the step which was found out in last process and the error η allowed by l, we could find out the height threshold ε_2 of the next step.

We assume $\varepsilon_2 = \varepsilon$. If the next step satisfies the step length support leg constraint of this movement type completely, and the length of its minimum support leg constraint frame is one integer in interval $[l - \eta, l + \eta]$, then ε_2 of this step is the same as ε(see Fig. 4).

If the next step does not satisfy the step length support leg constraint of this movement type, it means that ε_2 is too small (see Fig. 5). We need to add λ. And the size of λ should be suitable. If the size is too big, the length of minimum support leg constraint frame will bigger than $l + \eta$. If the size is too small, it may still can not satisfy the step length support leg constraint or maybe the length of

Fig. 3 We find the height threshold value ε of one step; *Top* the vertical trajectory of the ankle. *Bottom* characteristic phases of a motion

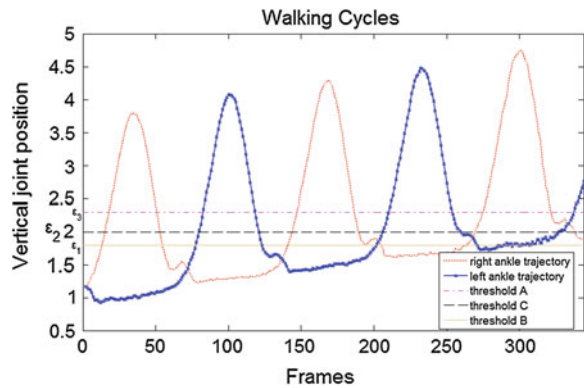**Fig. 4** Running movement: the height threshold of the first step and that of the second step is equal, both of them is ε
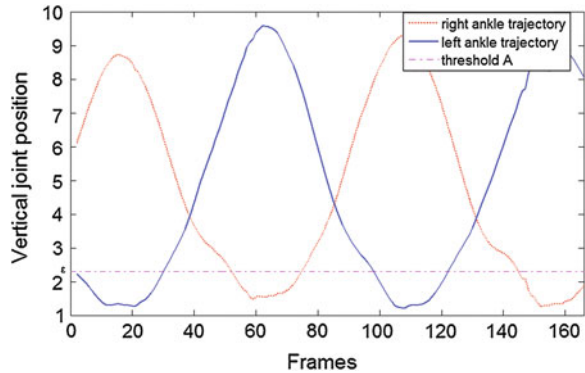
Fig. 5 Walking movement: the height threshold of the first step is ε_1, that of the second step is ε_2, and that of the third step is ε_3

minimum support leg constraint frame will smaller than $l - \eta$. We take $\lambda = 0.1$ in this example.

If the length of minimum support leg constraint frame is bigger than $l + \eta$, it means ε_2 is too big. We need to reduce λ.

If the length of minimum support leg constraint frame is smaller than $l - \eta$, it means ε_2 is too small. We need to increase λ.

(4) Repeat the third step until the final step.

5 Results

The testing results will produce a large error when we use a fixed size foot height threshold, because it is impossible that the data we captured have no noise. As shown in Fig. 5, if we use the unified size ε_1, the third step will not satisfy the step length support leg constraint and the support leg constraint of the second step would be too long or too short. If we use the unified size ε_3, the length of each step's support leg constraint will too long. If we adjust the movement direction in
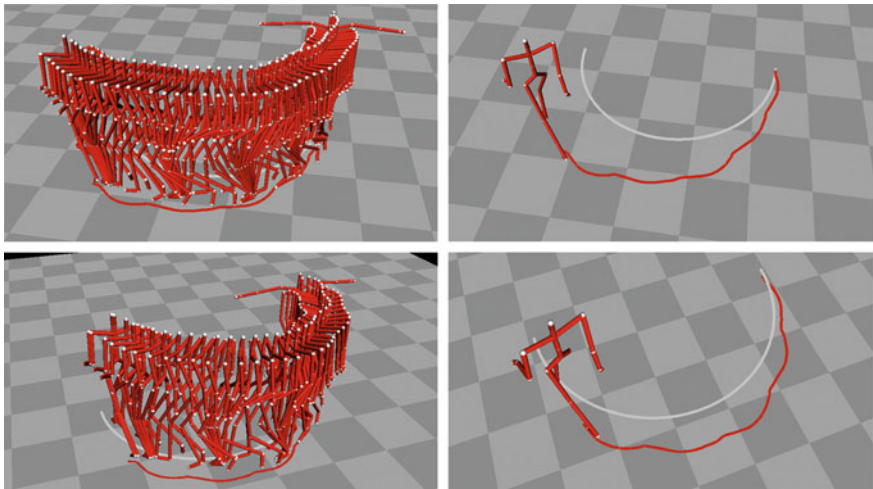
Fig. 6 Walk motion: using the unified ankle height threshold and adaptive method, compared with the edited circular path. *Up* unified; *Bottom* adaptive

A_{con} and N_{con} stage, slide phenomenon will appear. So we have to adjust the movement direction in L_{con} and R_{con} stage. It leads to big deviation between the edited motion path and the specified path, if A_{con} and N_{con} is too long. We could see from the Figs. 6, 7, 8 that, when we edit the motion [14] by foot height threshold adaptively method, the deviation between the edited motion path and the specified path is very small.
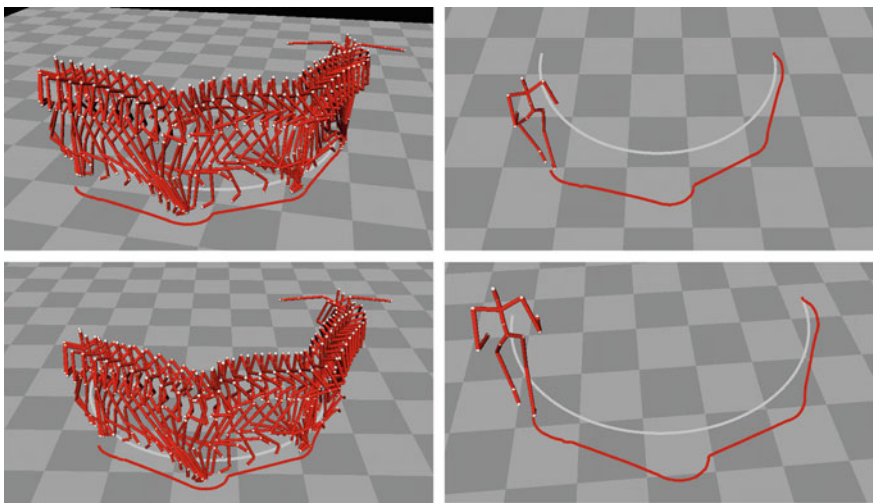
Fig. 7 Fast run motion: using the unified ankle height threshold and adaptive method, compared with the edited circular path. *Up* unified; *Bottom* adaptive

Fig. 8 Slow run motion: using the unified ankle height threshold and adaptive method, compared with the edited circular path. *Up* unified;*Bottom* adaptive
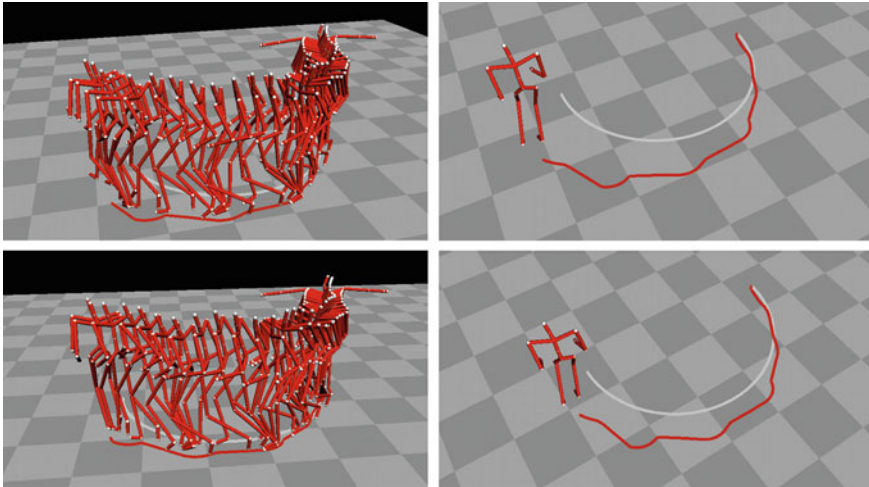
Table 1 Motion sequence

Name	Total frames	Step length
Walk	344 frames	3
Fast run	166 frames	2
Slow run	400 frames	4

This paper uses LUA + IUP + OpenGL codes to verify the availability of this algorithm. This paper uses three kinds of BVH data (see Table 1), which are downloaded from CMU database [13].

6 Conclusions and Future Work

Computer animation has formed a huge industry. Due to the faithful of the motion capture data, based on motion capture animation technology is gradually becoming a mature technology; Footplant detection plays a very important role on the effect of the movement editor or synthetic.

In this paper, we present a footplant detection method by adaptive determining foot height threshold. In our method, we firstly divide one step into a number of stages, each stage has different types of constraints. Then adaptively electing foot height threshold according to motion type and the length l of the minimum constraint frame that designed by the user and the maximum error φ of l to determine

which foot should be planted. Finally, we verify the effectiveness of our approach from the path offset error.

Our framework still has much space to be improved. For example, If the original movement data is standard without the noise point, then the threshold value of every step is same in our method, and the algorithm is more complex than the uniform threshold. Therefore, we have better effect on the noise point data processing. The threshold and error in the algorithm are defined by the user. On one hand, it increases the user interactivity. But on the other hand, it also increased the complexity of the algorithm. We are prepared to research how to determine those two values automatically in the future.

References

1. Lee J, Shin SY (1999) A hierarchical approach to interactive motion editing for human-like figures. In: Proceedings of ACM SIGGRAPH, Annual conference series, pp 39–48
2. Gleicher M (1999) Animation from observation: motion capture and motion editing. Compu Graph 51–55
3. Tolani D, Goswami A, Badler NI (2000) Real-time inverse kinematics techniques for anthropomorphic limbs. Graph Models 353–388
4. Kovar L, Schreiner J, Gleicher M (2002) Footskate cleanup for motion capture editing. In: Proceedings of the 2002 ACM symposium on computer animation (SCA)
5. Lyard E, Magnenat-Thalmann N (2007) A simple footskate removal method for virtual reality applications. Visual Comput 689–695
6. Guo XY, Xu SB, Che WJ, Zhang XP (2010) Automatic motion generation based on path editing from motion capture data. Trans Edutainment IVLNCS 6250:91–104
7. Gleicher M (1997) Motion editing with spacetime constraints. In: Proceedings of the 1997 symposium on interactive 3D graphics
8. Raunhardt D, Boulic R (2009) Motion constraint. Vis Comput 509–518
9. Liu CK, Popovic' Z (2002) Synthesis of complex dynamic character motion from simple animations. ACM Trans Graph 408–416
10. Glardon P, Boulic R, Thalmann D (2006) Robust on-line adaptive footplant detectionand enforcement for locomotion. Vis Comput 194–209
11. Gao Y, Ma LZ, Wu XM, Chen ZH (2005) Automatic foot-plant constraints detection shoes. In: Proceedings of the computer graphics, imaging and vision: new trends (CGIV'05)
12. Ikemoto L, Arikan O, Forsyth D (2006) Knowing when to put your foot down. In: Proceedings of the symposium on interactive 3D graphics and games, pp 49–53
13. http://mocap.cs.cmu.edu/info.php
14. Gleicher M (2001) Motion path editing. In: Proceedings of the 2001 symposium on Interactive 3D graphics, pp 195–202

Algorithmic Tile Self-Assembly for Solving the Maximal Matching Problem

Zhen Cheng, Yufang Huang and Jianhua Xiao

Abstract The maximal matching problem is a classic combinatorial optimization problem. Recently, computation by algorithmic tile self-assembly is proved to be a promising technique in nanotechnology, and this computational model is also demonstrated to be Turing universal. In this paper, the process of tile self-assembly model which is used to solve the maximal matching problem is shown including three operations: nondeterministic guess operation, AND operation and comparing operation. Our method can be performed this problem in $\Theta(mn)$ steps, here m and n is the number of edges and vertices of the given graph respectively.

Keywords Algorithmic · Tile · Self-assembly · Maximal matching problem

1 Introduction

In 1994, Adleman demonstrated that the DNA recombination techniques can be used to solve combinational search problem [1]. Compared to conventional silicon-based computers, DNA computing potentially provides more parallelism and higher density storage [2]. In recent years, many models and algorithms of DNA computing are extensively investigated [3]. Algorithmic tile self-assembly is an important model of DNA computing which is studied by many researchers.

Z. Cheng (✉)
College of Computer Science and Technology, Zhejiang University of Technology,
Hangzhou 310023, China
e-mail: chengzhen0716@163.com

Y. Huang
College of Mathematics, Southwest Jiaotong University, Chengdu 610031, China
e-mail: huangyufang@home.swjtu.edu.cn

J. Xiao
The Research Center of Logistics Nankai University, Tianjin 300071, China
e-mail: jhxiao@nankai.edu.cn

Seeman proposed DNA nanotechnology to make self-assembled nanostructures from DNA molecules [4]. Based on this landmark work, Seeman realized a patterned lattice and the complex algorithmic pattern [5]. Barish et al. [6] experimentally performed patterns of binary counters using implementation of copying and counting. The 3D nanostructures [7] can also be created. The abstract Tile Assembly Model [8] is a formal model of crystal growth which is designed to model self-assembly of molecules. Rothemund [9] defined the abstract tile assembly model which extended the theory of Wang tilings [10]. Murata [11] proposed a precise simulation model for tile self-assembly. Algorithmic tile self-assembly is demonstrated as a promising technique for solving NP-complete problems with high probability [12]. The complexity of self-assembly nanostructure is also studied [13]. Much attention has been paid to algorithms which can be used in research of solving the maximal matching problem [14]. Yin [15] proposed a DNA algorithm for maximal weight matching problem. On these basis, an important method of DNA computing which is algorithmic tile self-assembly model is used to implement the maximal matching problem in $\Theta(mn)$ steps, here n is number of the vertices and m is the number of the edges of the given graph.

The rest of the paper is structured as follows: Sect. 2 will show how the maximal matching problem is solved based on this model and give the complexity analysis of this algorithm, followed by some conclusions in Sect. 3.

2 Implementing the Maximal Matching Problem

In this section, the basic tile units used in this algorithmic tile self-assembly model are constructed, then the architectures of success and failure in tile attachments are given to demonstrate the reasonability and validity of this model.

Mathematically, given an undirected graph $G = (V, E)$, suppose the number of the vertices and edges of this graph be n and m respectively, a matching in a graph G is a subset M of the edges of G such that no two distinct edges in M are adjacent to the same vertex. Here the maximal matching problem is that of finding a maximum-cardinality subset M of the edges set E in this graph G. The edge e_i is in the matching set which can be represented as $e_i = 1$ otherwise $e_i = 0$. Thus the number of edges in the matching set can be expressed by $Z = \Sigma e_i$. Thus this problem can be mathematically transformed into the optimization integer programming problem:

$$Max\, Z = \sum e_i$$

Subject to

$$\begin{cases} e_i + e_j \le 1 & (i \ne j) \\ e_i \in \{0,\ 1\} & \forall e_i \in E \end{cases}$$

Fig. 1 The structure of basic
tile unit

According to the definition of the maximal matching problem, the solution of
the matching problem is equal to 0–1 programming problem. The coefficient
matrix of the inequalities can be represented a matrix $C = (c_{ij})$, here the element
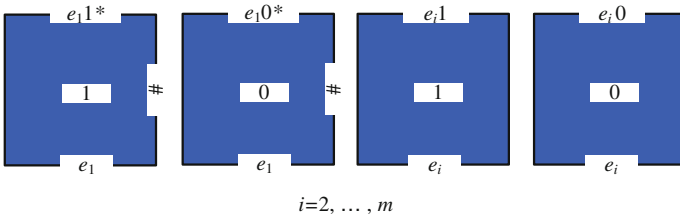c_{ij} of this matrix is 0 or 1.

2.1 Designing the Basic Tile Units

Here, the abstract TAO tile assembly model is mainly used to implement the
maximal matching problem which is shown in Fig. 1. Intuitively, the right square
is used to abstract the structure of the left tile. The binding domains sets including
σ_N, σ_E, σ_S and σ_W are on its four sides.

Given an undirected graph $G = (V, E)$, suppose the number of the vertices and
edges of this graph be n and m respectively. The adjacent matrix of this graph can
be obtained, and the inequalities erected on the relationship of the edges are as
above. In order to check whether the value of the edges are satisfied with the
inequalities, the value of the edges should be firstly nondeterministically guessed
which is 0 or 1. The binding domains set of this nondeterministic guess operation
is $\sum_{guess} = \{\#, e_1, e_i, e_i 1, e_1 0, e_i 1, e_i 0\}$ and the tile set T_1 is as follows in Fig. 2.

The first two tiles of the tile types in Fig. 2 can nondeterministically generate
the value of the first edge e_1 which is 0 or 1. And this value can be passed to upper
from the bottom of this tile with the label "$e_1 0^*$" or "$e_1 1^*$". The last two tiles in
Fig. 2 can nondeterministically guess the value of the edges e_2 to e_m. As the same
representation, the value is passed to upper from the bottom of this tile with the
label "$e_i 0$" or "$e_i 1$". Here, the label "$*$" is used to distinguish the edge e_1 from the
other edges which should be firstly made AND operation at the next step.

Once the value of the edges are obtained by the nondeterministic guess operation,
it is made AND operation with the coefficients in the coefficient matrix $C = (c_{ij})$. By

$i=2, \ldots, m$

Fig. 2 The tile set T_1 of nondeterministic guess operation can generate the value of the edges,
which is 0 or 1
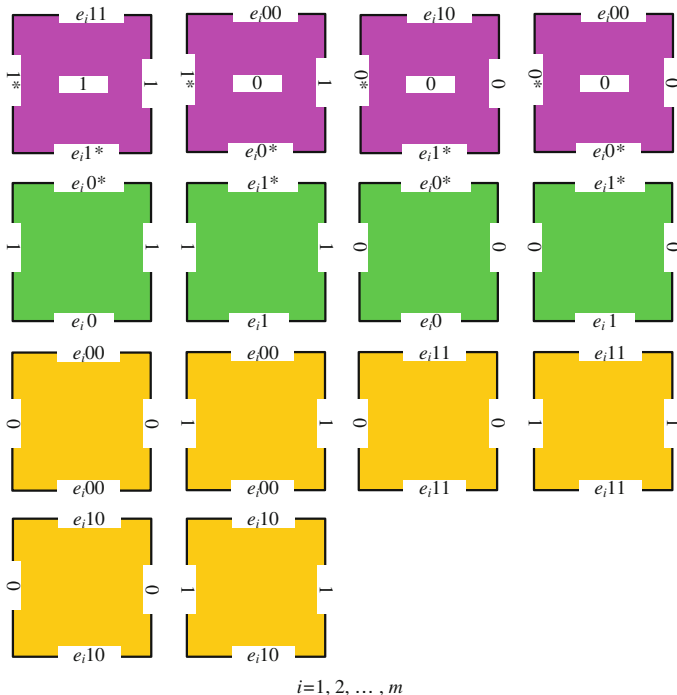
Fig. 3 The tile set of AND operation

this matrix, it is shown which edges are adjacent in the graph. The binding domains set of AND operation is $\sum_{AND} = \{1, 1^*, 0, 0^*, e_i1^*, e_i0^*, e_i11, e_i10, e_i00, e_i1, e_i0\}$ and the basic tile set T_2 are shown in Fig. 3.

The AND operation is made between the value of edges and the corresponding coefficients of edges with the start label "*". The tiles in the first row show how the AND operation is carried out. The bottom of the tile is "e_i1^*" or "e_i0^*", and it means value of the edge e_i is 1 or 0. The right of the tile is 1 or 0 which represents the coefficient of the edge e_i, so the middle of the tile is the AND operation result. The value of the edge e_i is passed to the upper and the coefficient of the edge e_i should be passed to the left of the tile. The result of the AND operation is 1 if two input numbers are two ones, but it is 0 if one of the two numbers is 0.

The tiles in the second row in Fig. 3 show the value of the edges which should be passed to the upper layer to make the AND operation for these edges. The inputs including the value of the bottom and the right should be passed to the upper and the left respectively. The bottom of the tiles in the third and forth row is "e_i10", "e_i11" or "e_i00" which means the value of the edge e_i is 1, 1, 0 and the result of AND operation from the previous step is 0, 1, 0 respectively. It is passed to the upper no matter what the input of the right is 0 or 1.

In order to check whether the values of the edges are satisfied with the inequalities, the comparing operation is carried out. If the inequalities are
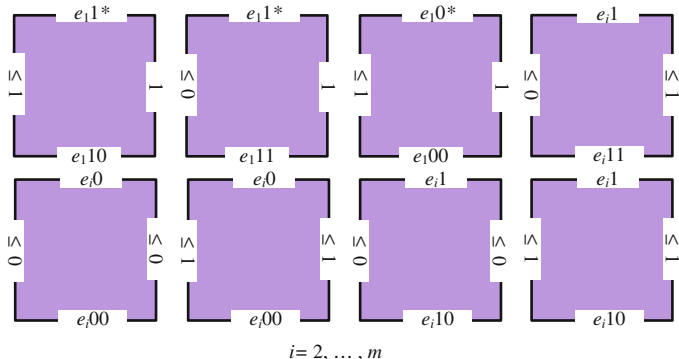
e_11^* e_11^* e_10^* e_i1

≥ 1 1 ≥ 0 1 ≥ 1 1 ≥ 0 1

e_110 e_111 e_100 e_i11

e_i0 e_i0 e_i1 e_i1

≥ 0 ≥ 0 ≥ 1 ≥ 1 ≥ 0 ≥ 1 ≥ 0 1

e_i00 e_i00 e_i10 e_i10

$i = 2, \ldots, m$

Fig. 4 The basic tile set of the comparing operation

demonstrated to be satisfied, it shows that the edges in these inequalities are adjacent. The right of the inequalities is 1, so the upper bound of the left is 1, and it also shows that no more than two edges are adjacent in the inequalities. The binding domains set of this comparing operation is $\sum_{comparing} = \{1, \ \leq 1, \ \leq 0, e_11^*, e_10^*, e_i11, e_i00, e_i10, e_i00, e_i1, e_i0\}$ and the tile set T_3 is as follows in Fig. 4.

The right of the first three tiles in the first row in Fig. 4 is 1 which is the upper bound of the inequality, and the bottom of the tile is "e_110", "e_111" or "e_100" which means the value of the edge e_1 is 1, 1, 0 and the result of the AND operation from the previous step is 0, 1, 0 respectively. The value of e_1 can be passed to the upper of the tile. If the AND operation result is 1, the subtraction can be made between 1 and the AND operation result, so the relationship result can be represented as "≤ 0", here 0 is the subtraction result and "$\leq$" means the value of the left is no more than the right of this inequality. If the AND operation result is 0, the relationship result is "≤ 1", here 1 is the subtraction result. For the forth tile in the first row, the inputs are "≤ 1" and "e_i11" as the right and bottom of this tile respectively. The value of the edges is passed to the upper of the tile.

For the tiles in the second row in Fig. 4, the relationship result of is "≤ 0" or "≤ 1", and the bottom of the tile as the input is "e_i10" or "e_i00", then the process of the comparing is operated as the forth tile in the first row.

The boundary tiles include input and output tiles, and computation boundary tiles which are shown as following in Fig. 5.

2.2 Architectures for Solving the Maximal Matching Problem

Now an example of the maximal matching problem is taken to show the process of the method based on algorithmic tile self-assembly. Given an undirected graph $G = (V, E)$ as follows in Fig. 6 as follows.
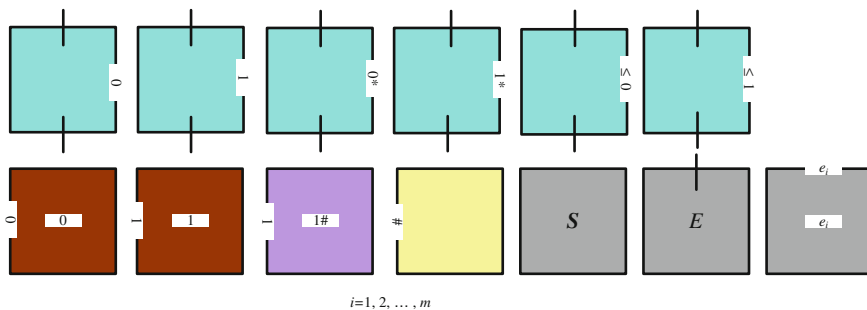
Fig. 5 The basic tile types of boundary tiles

Fig. 6 Given an undirected
graph $G = (V, E)$, where
$V = \{v_1, v_2, v_3, v_4, v_5, v_6\}$,
$E = \{e_1, e_2, e_3, e_4, e_5\}$
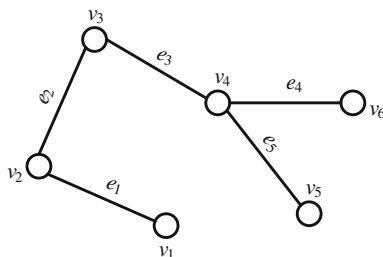
First, according to the adjacent matrix of this graph, this problem is mathematically transformed into the following 0–1 programming problem as follows. Here, the value of the edges satisfies with $e_i \in \{0, 1\}$ $(i = 1, 2, 3, 4, 5)$. The vector $M = (e_1, e_2, e_3, e_4, e_5)$ is used to denote a matching set.

$$
\begin{cases}
e_1 + e_2 \leq 1 \\
e_2 + e_3 \leq 1 \\
e_3 + e_4 \leq 1 \\
e_3 + e_5 \leq 1 \\
e_4 + e_5 \leq 1
\end{cases}
$$

Second, when all the tiles and the seed configuration are prepared, they are put together into the reaction buffer. According to the mechanism of algorithmic tile self-assembly through Watson–Crick base pairing, the self-assembly process starts at the same time with the connector tiles, and the final stage is seen in Fig. 7.

The nondeterministic guess operation generates the random value of the edges. In this example, the value of the edges are given as $M = (0, 1, 0, 1, 0)$. The value of the edge is labeled as $e_i \in \{0, 1\}$ $(i = 1, 2, 3, 4, 5)$. The first inequality is composed of the edges e_1 and e_2, and the AND operation is made between e_i and the corresponding coefficient of e_i, here the AND result is 0, 1, 0, 0, 0 for the five edges e_i respectively. Then the comparing operation can give the relationship of the right and left of the inequality. The AND operation result for the first edge e_1

Fig. 7 The final stage of the successful example
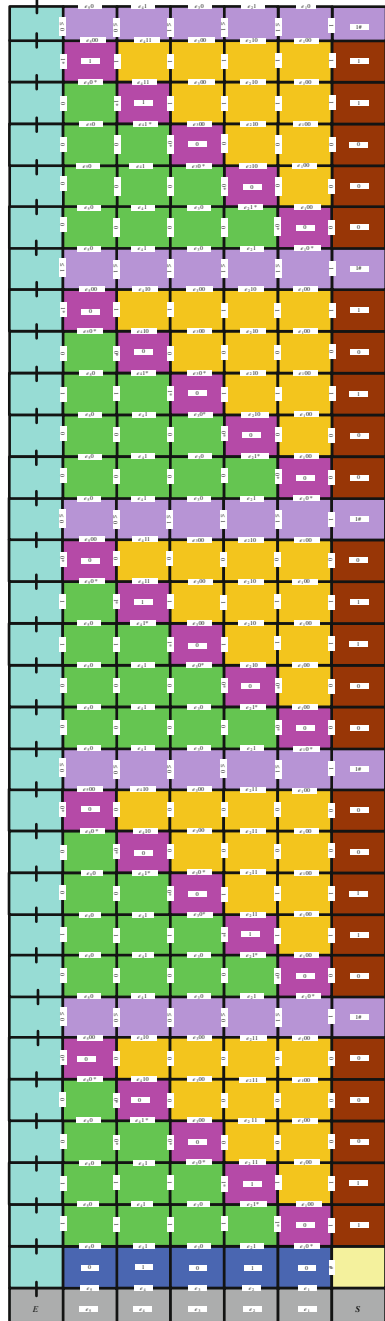
and the corresponding coefficient is 0 which should be made the subtraction operation and the comparing result is " ≤ 1 ". The AND operation result for the second edge e_2 and the corresponding coefficient is 1 with the input from the right of the tile " ≤ 0 ", so the comparing result is " ≤ 0 " at this step. The value of the five edges e_i ($i = 1, 2, 3, 4, 5$) is satisfied with this inequality.

Thus the other four inequalities are also checked by the same process using the AND operation and comparing operation. The vector $M = (0, 1, 0, 1, 0)$ satisfies the other four inequalities. Finally, to output the result of computation, a modification of the standard sequence-reading operation that uses a combination of PCR and gel electrophoresis is implemented. Through using the operations, the strands of different lengths representing the output tiles in the result strand are extracted. In this example, $M = (0, 1, 0, 1, 0)$ is the feasible solution of this problem.

For the same matching problem, the nondeterministic algorithm which shows the failure in attaching tiles is also given, and it doesn't get the feasible result in Fig. 8. Firstly, the nondeterministic guess operation generates one of random value of the edges which is represented as $M = (1, 1, 0, 0, 0)$. The process of the growth for the assembly complexes is the same as the former example. The AND operation result is 1, 1, 0, 0, 0 for the five edges e_i ($i = 1, 2, 3, 4, 5$) respectively in the first inequality. The values of the edges are checked that they can't be satisfied with the first inequality by the comparing operation. The result is "No tile can match", thus the value of the edges is not the feasible solution of this problem.

For a maximal matching problem, in terms of the complexity of the design for the maximal matching problem using tile self-assembly model, here the computation time and number of distinct tiles required are mainly considered. The computation time means the assembled steps which are needed to complete the growth of the complexes, and it is equal to $(n + 1) + m(n + 1) = mn + m+n + 1 = \Theta(mn)$. Here, the number of the vertices and edges of this graph is n and m respectively.

The basic tiles needed in the process of performing this problem contain the computation tiles, boundary tiles. For the nondeterministic guess operation in Fig. 2, there needs $2m$ tile types. For the tile types used in the AND operation in Fig. 3, the number of the tile types for this operation is $11m$. For the comparing operation in Fig. 4, the number of the tile types for the comparing operation is $(5m - 2)$. The basic tile types of boundary tiles are seen in Fig. 5, so the number of the boundary tiles is $(m + 12)$. The total number of tiles is $19m + 10 = \Theta(m)$. Therefore the maximal matching problem is carried out with $\Theta(m)$ tiles in $\Theta(mn)$ steps.

Fig. 8 A failure example for the maximal matching problem
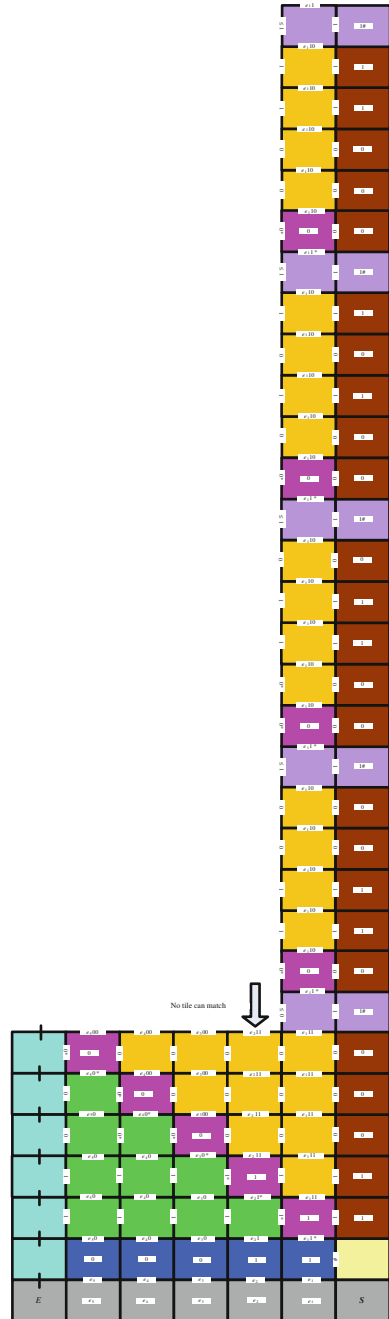
3 Conclusions

The algorithm for implementing the maximal matching problem using tile self-assembly model is mainly proposed. First, the nondeterministic guess operation generates the value of the edges in the given graph, then the edges from e_1 to e_m and the corresponding coefficients can be made by AND operation, thus the value of the edges are checked whether they are satisfied with each inequality by the comparing operation. To conclude, this model can be successfully performed the maximal matching problem in $\Theta(mn)$ steps which is linearly to the multiplication of the number of the vertices and edges in the given graph, once the initial strands are constructed. The field of algorithmic tile self-assembly holds tremendous promise, and it also might have other applications in biology and other ways.

References

1. Adleman LM (1994) Molecular computation of solutions to combinatorial problems. Science 266:1021–1024
2. Pan LQ, Liu GW, Xu J (2004) Solid phase based DNA solution of the coloring problem. Prog Nat Sci 14:104–107
3. Pan LQ, Pérez-Jiménez MJ (2010) Computational complexity of tissue-like P systems. J Complex 26:296–315
4. Seeman NC (1998) DNA nanotechnology: novel DNA constructions. Annu Rev Biophy Biomol Struct 27:225–248
5. Mao C, Sun W, Seeman NC (1999) Designed two dimensional DNA Holliday junction arrays visualized by atomic force microscopy. J Am Chem Soc 121:5437–5443
6. Barish R, Rothemund PW, Winfree E (2005) Two computational primitives for algorithmic self-assembly: copying and counting. Nano Lett 12:2586–2592
7. Lo PK, Metera KL, Sleiman HF (2010) Self-assembly of three-dimensional DNA nanostructures and potential biological applications. Curr Opin Chem Biol 14:597–607
8. Wang H (1961) Proving theorems by pattern recognition. I. Bell Syst Tech J 40:1–42
9. Winfree E (1998) Algorithmic self-assembly of DNA. Ph.D. thesis, Caltech, Pasadena, CA
10. Rothemund P, Winfree E (2000) The program-size complexity of self-assembled squares. In: Proceedings of the annual ACM symposium on theory of computing, Portland, USA, pp 459–468, May 21–23
11. Fujibayashi K, Murata S (2009) Precise simulation model for DNA tile self-assembly. IEEE T Nanotechnol 8:361–368
12. Brun Y, Reishus D (2009) Path finding in the tile assembly model. Theor Comput Sci 410:1461–1472
13. Reif JH, Sahu S, Yin P (2011) Complexity of graph self-assembly in accretive systems and self-destructible systems. Theor Comput Sci 412:1592–1605
14. Cardinal J, Labbe M, Langerman S et al (2005) A tight analysis of the maximal matching heuristic. Lect Notes Comput Sci 3595:701–709
15. Ding X, Yin ZX, Wang Y (2010) A DNA algorithm for maximal matching problem. International conference on bio-inspired computing: theories and applications, Beijing, China, pp 143–147

Application of FDTD and Harmony Search Algorithm for Inverse Scattering Problem

Jing Lu, Junhua Gu and Yang Lu

Abstract Electromagnetic inverse scattering of two-dimensional (2-D) perfectly conducting objects with transverse magnetic (TM) wave incidence by the harmony search algorithm (HS) is presented. The idea is to perform the image reconstruction by utilisation of HS to minimise the discrepancy between the measured and calculated scattered field data. Finite difference time domain method (FDTD) is used to solve the scattering electromagnetic wave of the objects. The efficiency of applying the above methods for microwave imaging of a two-dimensional perfectly conducting cylinder are examined. Numerical results show that good reconstruction can be obtained by the optimisation method.

Keywords Electromagnetic inverse scattering · Harmony search algorithm · FDTD

1 Introduction

Inverse scattering problems, have got much interest in the past few years, involve determining the location, shape and material parameters of the unknown objects from the measured fields that are scattered by the objects, when illuminated by known incident fields. Numerical inverse scattering studies are based on either frequency or time domain approaches. Time-domain inverse-scattering problems somewhat related to the present study commonly appear in the area of geosciences

J. Lu (✉)
School of Electrical Engineering, Hebei University of Technology, Tianjin 300401, China
e-mail: lujing@scse.hebut.edu.cn

J. Lu · J. Gu · Y. Lu
School of Computer Science and Engineering, Hebei University of Technology,
Tianjin 300401, China
e-mail: jhg@scse.hebut.edu.cn

and remote sensing [1–3]. The scatterer reconstruction belongs to the general category of limited angle microwave imaging problems. These problems are known to have nonlinear and ill-posed properties due to the lack of the measured information and the multi-scattering effects between the objects. In general, the ill-posedness can be treated by traditional regularization schemes or transformed into a optimization problem. An error or objective function is defined as a function of the difference between the measured and calculated field, we use the optimization algorithm to reconstruct the unknown objects by minimizing the objective function [4, 5]. In general, the nonlinearity of the problem is coped with by applying iterative optimization techniques. Those algorithms used stochastic strategies, which are global searching algorithm, compared with local inversion algorithm, they have some advantages including robustness and strong searchability simplicity.

Some population-based stochastic methods are used to search the global extreme of the inverse problems, such as genetic algorithm (GA), particle swarm optimization (PSO).

To this day, there is still no literature using the harmony search algorithm (HS) to reconstruct the electromagnetic imaging of perfectly conducting cylinder with circular shape in free space under time domain.

In this paper, the FDTD method is used for the forward problem at each iteration step, which calculates the electric and magnetic fields using the coupled Maxwell's curl equations. To reduce the computational domain and nonphysical reflections at the outer boundaries, an absorbing boundary condition (ABC) is required. In this paper, the perfectly matched layer (PML) method which was introduced by Berenger is used.

2 Forward Problem

Let us consider the problem geometry of microwave imaging shown in Fig. 1. The cylinder is parallel to z axis and located within an investigation domain S, while the cross-section of the cylinder is circular. A uniform circular array of microwave transceivers surround the investigation domain S. The object is illuminated by a Gaussian pulse line source located at the points denoted by Tx and reflected waves are recorded at those points denoted by Rx. By looking at the x-y plane, the PEC cylinder has a circular cross-sectional contour characterised by the center position (x0,y0) and the radius r. The computational domain is discretised by Yee cells. The forward problem is implemented in the FDTD code [6]. To reduce the computational domain and the reflection from the perfect matching layer interface, an absorbing boundary condition is required. In this paper, the perfectly matched layer (PML) method is used.

Fig. 1 Geometrical
configuration of the problem
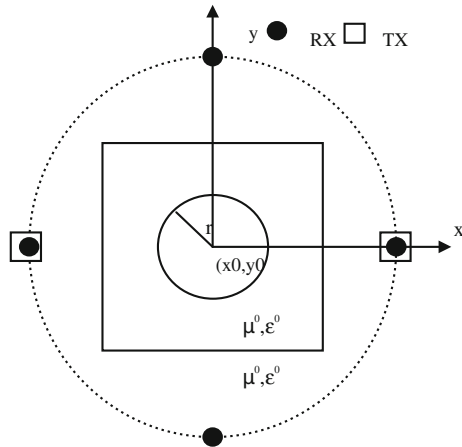
3 Inverse Problem

For the inverse scattering problem, the shape and location of the perfectly con-
ducting cylinder are reconstructed by the given scattered electric field obtained at
the receivers. This problem is resolved by an optimization approach, for which the
global searching scheme harmony search algorithm (HS) is employed to minimize
the following cost function (CF):

$$CF = \frac{\sum\limits_{n=1}^{N}\sum\limits_{m=1}^{M}\sum\limits_{q=0}^{Q}\left|E_z^{\exp}(n,m,q\Delta t) - E_z^{cal}(n,m,q\Delta t)\right|}{\sum\limits_{n=1}^{N}\sum\limits_{m=1}^{M}\sum\limits_{q=0}^{Q}\left|E_z^{\exp}(n,m,q\Delta t)\right|}$$

where $E_z^{\exp}$ and E_z^{cal} are experimental electric fields and the calculated electric
fields, respectively. The N and M are the total number of the transmitters and
receivers, respectively. Q is the total time step number of the recorded electric
fields.

4 Harmony Search Algorithm

HS is a recently-developed meta-heuristic optimization algorithm mimicking
music improvisation, where music players improvise the pitches of their instru-
ments to obtain better harmony [1]. As a meta-heuristic algorithm, HS is similar to
other meta-heuristic algorithms such as genetic algorithm and particle swarm
algorithm because they manage a group of solutions in a population rather than a
single solution with gradient information. However, HS has its own uniqueness.
While genetic algorithm generates a new offspring from two parents in the

population, HS generates it from all of the existing vectors stored in HM [7–9]. The solution of an optimization problem using HS is performed based on the following five steps:

Step 1. Initialize the problem and algorithm parameters.

In Step 1, the optimization problem is specified as follows:

Minimize f(x) subject to $x_i \in X_i$, $i = 1,2,…,N$, where f(x) is an objective function; x is the set of each decision variable x_i; N is the number of decision variables, X_i is the set of the possible range of values for each decision variable, that is $Lx_i \le X_i \le Ux_i$, and Lx_i and Ux_i are the lower and upper bounds for each decision variable. The HS algorithm parameters are also specified in this step. These are the harmony memory size (HMS), harmony memory considering rate (HMCR); pitch adjusting rate (PAR); and the number of improvisations (NI), or stopping criterion.

Step 2. Initialize the harmony memory.

The HM matrix filled with as many randomly generated solution vectors as the HMS, is as follows:

$$\begin{pmatrix} x_1^1 & x_2^1 & x_{N-1}^1 & \cdots & x_N^1 & f(\vec{x}^1) \\ x_1^2 & x_2^2 & x_{N-1}^2 & \cdots & x_N^2 & f(\vec{x}^2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1^{HMS-1} & x_2^{HMS-1} & x_{N-1}^{HMS-1} & \cdots & x_N^{HMS-1} & f(\vec{x}^{HMS-1}) \\ x_1^{HMS} & x_2^{HMS} & x_{N-1}^{HMS} & \cdots & x_N^{HMS} & f(\vec{x}^{HMS}) \end{pmatrix}$$

In this step, a new solution vector $\vec{x}' = (x_1', x_2', x_3', …, x_N')$ is generated based on the three rules (memory consideration, pitch adjusting, and random selection).

In the memory consideration, the value of a decision variable $(x_i',\ 1 \le i \le N)$ for the new vector is chosen from any of the values in the specified HM range $(x_i^1 \sim x_i^{HMS})$. The HMCR, which varies between 0 and 1, is the rate of choosing one value from the historical values stored in the HM, whereas $(1 - HMCR)$ is the probability of randomly selecting from the possible range.

$$x_i' \leftarrow \begin{cases} x_i' \in \{x_i^1, x_i^2, …, x_i^{HMS}\} & \text{with probability HMCR,} \\ x_i' \in X_i & \text{with probability (1-HMCR).} \end{cases}$$

After performing the selection procedure, each decision variable is evaluated to determine if the pitch adjusting is necessary or not. This operation uses the PAR parameter, which is the rate of pitch adjustment as follows:

If rand(0,1) < PAR
$x_i' \leftarrow x_i' \pm rand(0, 1) \times bw$
Else
$x_i' \leftarrow x_i'$
End if

where bw is an arbitrary distance bandwidth, rand ( ) is a random number between 0 and 1, and PAR is pitch adjusting rate. The traditional HS algorithm uses fixed value for PAR and bw, IHS algorithm [10] proposed by M. Mahdavi uses variables PAR and bw. PAR is a function of the number of iterations (currentIteration). The function is as follows:

PAR=(PARmax-PARmin)/(MaxItr)*currentIteration + PARmin

PARmax and PARmin is the maximum value and the minimum value of PAR. MaxItr is the maximum number of iterations. In this step, HM consideration, pitch adjustment or random selection is applied to each variable of the new harmony vector in turn.

Step 3. Update the harmony memory.

In this step, the new improvised harmony $\vec{x}' = (x_1', x_2', x_3', \ldots, x_N')$ and worst harmony in HM is compared in terms of their objective function values. If the new harmony vector is better than the worst harmony in the HM, the new harmony is included in the HM and the existing worst harmony is taken out.

Step 4. Check the stopping criterion.

The optimization algorithm continues to be executed by iterating from Step 3 to 4 until the stopping criterion (maximum number of improvisations) are satisfied.

5 Numerical Results

As shown in Fig. 1, the problem space is divided in 279×155 grids with the grid size $\Delta x = \Delta y = 5.95$ mm. The metallic cylinder is located in free space. The cylindrical object is illuminated by a transmitter at two different positions, N = 2. The scattered E fields for each illumination are collected at the four receivers, M = 4. The excitation source is the Gauss pulse. The time duration is set to $224\Delta t$. There are three unknown parameters to retrieve, which include the center position(x0,y0), the radius r. Searching ranges used for the HS to optimize the cost function CF are [r-10, r + 10], [x0-10, x0 + 10], [y0-10, y0 + 10]. The parameters are set as below: HMCR is set to 0.9, HMS is set to 6, PARmin is set to 0.4, PARmax is set to 0.9, bw is set to between 0.0001 and 1.0. The largest number of iterations is set to 2,000. Reconstruction of a perfectly conducting circular cylinder by the HS is shown in Fig. 2.
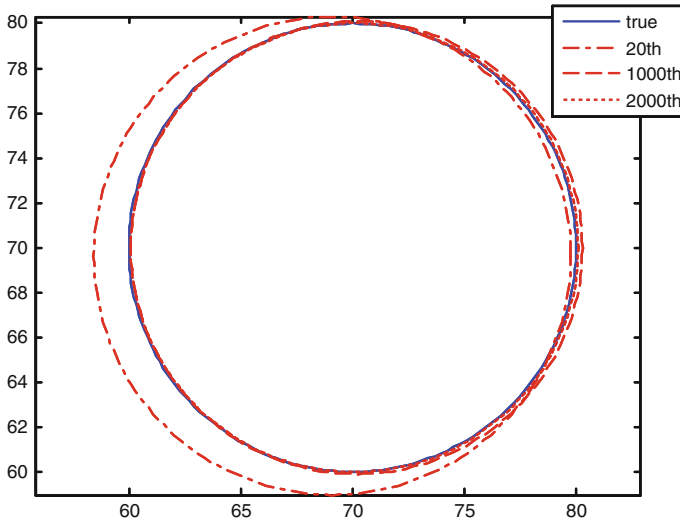
Fig. 2 Reconstruction of a perfectly conducting circular cylinder by the HS. The radius of the cylinder is 20, and the center is (70,70)

6 Conclusion

In this chapter, we present a method of applying harmony search algorithm to reconstruct the shape of a perfectly conducting circular cylinder by TM waves. By combining the FDTD method and the harmony search algorithm, good reconstructed results are obtained. The inverse problem is reformulated into an optimization one, and then the harmony search algorithm is employed to search the parameter space. It converges to the global extreme, while the calculus-based method often becomes trapped in a local extreme. From our experience, the main difficulties in applying the harmony search algorithm to this problem are how to choose the parameters, such as harmony memory considering rate and the number of improvisations. Different parameter sets will affect the speed of convergence and the computing time required.

References

1. Geem ZW, Kim JH, Loganathan GV (2001) A new heuristic optimization algorithm: harmony search. Simulation 76(2):60–68
2. Chew WC, Otto GP (1992) Microwave imaging of multiple conducting cylinders using local shape functions. IEEE Microwave Guided Wave Lett 2:84–286

3. Takenaka T et al (1997) Local shape function combined with genetic algorithm applied to inverse scattering for strips. Microwave Opt Tech Lett 16(6):337–341
4. Qing A (2001) An experimental study on electromagnetic inverse scattering of a perfectly conducting cylinder by using the real-coded genetic algorithm. Microwave Opt Technol Lett 30:315–320
5. Sun CH, Chiu CC, Li CL (2010) Time-domain inverse scattering of a two-dimensional metallic cylinder in slab medium using asynchronous particle swarm optimization. Prog Electromagn Res M. PIER M, 14:85–100
6. Abenius E, Strand B (2006) Solving inverse electromagnetic problems using FDTD and gradient-based minimization. Int J Numer Meth Eng 68(6):650–673
7. Geem ZW (2009) Music-inspired harmony search algorithms: theory and applications. Verlag, Germany
8. Geem ZW (2009) Harmony search algorithms for structural design optimization. Springer, Germany
9. Lee KS, Geem ZW (2004) A new structural optimization method based on the harmony search algorithm. Comput Struct 82(9–10):781–798
10. Mahdavi M (2007) An improved harmony search algorithm for solving optimization problems. Appl Math Comput 188:1567–1579

Simulation of Space Information Network SCPS-TP Based on OPNET

Yufei Zhang, Jinhai Su and Chuanfu Zhang

Abstract In order to prove the performance of SCPS-TP which is proposed by CCSDS in researching space information network, this paper designs a SCPS-TP model including SNACK module, congestion control module and header compression module based on the three key techniques of SCPS-TP using OPNET. A space information network simulation platform is built, and the performance of SCPS-TP is simulated and compared to TCP. The simulation results show that SCPS-TP performs better in high Bit Error Rate and long round-trip delay environment than TCP.

Keywords SNACK · Vegas · Header compression · OPNET · Protocol simulation

1 Introduction

Space information network is composed of space, near space, air, ground network. Space-based network which is mainly composed of multiple orbit satellite is the core of space information network. Relative to the ground network, the space information network has longer transmission time, greater delay bandwidth product, higher error rate and intermittent outage characteristics [1]. If the transport layer control protocol TCP which is widely used in the ground network

Y. Zhang (✉) · J. Su · C. Zhang
Zhengzhou Information and Science Technology Institute, Zhengzhou, China
e-mail: zyf0712@sina.cn

J. Su
e-mail: sujinhai@sina.com

C. Zhang
e-mail: zhangchuanfu@hotmail.com

directly applies in space information network, there will be problems concluding connection establishes round-trip time being too long, slow start spending too much time, packet loss causes judgment being not accurate, and retransmission strategy doesn't fit. These problems severely degrade the performance of TCP, and can't even ensure normal valid data transmission [2].

In view of this situation, the Consultative Committee for Space Date System (CCSDS) developed Space Communications Protocol Specification-transport layer protocol (SCPS-TP) in 1999. SCPS-TP has been widely used in space information network, but has also exposed some problems to be solved. In 2006 SCPS-TP were revised, being more suitable for space information network [3]. SCPS-TP protocol is based on TCP protocol, and is the extensions and modifications of TCP protocol [4].

This paper studies the SCPS-TP protocol. In the paper, a space information network simulation platform is established based on OPNET, and the protocol model is realized. Through simulation model the transmission performance of SCPS-TP with TCP in space information network is analyzed and verified at different bit error rate.

2 SCPS-TP Protocol

In the CCSDS SCPS protocol cluster, SCPS-TP is the key to achieve reliable data transmission. It consists of standard TCP protocol revision and a series of extended terms and enhanced items. Some expansion items are modification for standard TCP, and some other extended terms ensure that it won't affect network system interoperability [5]. SCPS-TP is based for communication between different transport layer protocols to take adaptive conversion [6]. The major improvement of SCPS-TP is reflected in three aspects: Selective negative acknowledgment (SNACK), congestion control, and header compression.

2.1 SNACK Options

Different from the ground network, there are three main reasons according to the loss of data in the space information network: link congestion, link corruption and link outage. SCPS-TP is able to identify different data loss reasons and take different measures [7]. The main reason for the loss of space information network data is the link corruption. SNACK is used to measure link corruption for SCPS-TP [8]. SNACK integrates Selective acknowledgement (SACK) and Negative acknowledgement (NAK) thought. Traditional TCP can confirm the receiving end of an error, while SNACK can mark the receiving end multiple errors and does not change the ACK. It instantly passed the ACK number to speed up recovery and avoid the transmission side translate into the window limited state. When the loss of the data segment appears, it allows both parties in communication to precede the

communication. The measure to avoid entering a congestion control can improve the link utilization and the recovery rate of the missing data segments [9].

2.2 Congestion Controls

SCPS-TP has two alternative congestion control algorithms: Van Jacobson's algorithm (VJ) and Vegas [10]. This paper takes Vegas as the research object. Different from VJ algorithm of TCP, which determines congestion happens through packet loss, Vegas use the Round-Trip Time (RTT) changes triggered by the desired throughput and the actual throughput difference to determine whether congestion occurs. It is possibly more accurate in predicting the link utilization, with better fairness and efficiency. It is better to use Vegas to avoid congestion, and capture the corruption caused by network link packet lost to congestion. Vegas have slow improvement to increase a trigger mechanism *diff* between the exponential growth phase and the linear growth stage. We define two thresholds α and β $(\alpha < \beta)$ to denote the redundant data network too little or too much. The adjusting method of the congestion window is shown below:

$$cwnd = \begin{cases} cwnd + 1 & diff < \alpha \\ cwnd & \alpha < diff < \beta \\ cwnd - 1 & diff > \beta \end{cases}$$

Compare *diff* with α and β, it is taken α as 1, β as 3. If $diff < \alpha$, the network congestion state is considered not serious, the congestion window can be further expanded to increase the transmission rate. If $\alpha < diff < \beta$, the network congestion state is steady, congestion window can be hold unchanged. If $diff > \beta$, network congestion is serious, the congestion window is reduced to decrease the transmission rate to avoid congestion occurs.

2.3 Header Compression

SCPS-TP uses end-to-end header compression technology [11] to solve the problem of limited bandwidth in the space information network. It is different from the TCP/IP header compression scheme which is placed in the link layer connection state table and based on hop-by-hop compression method. It can tolerate the loss of data segments and disconnection. A packet loss does not affect the subsequent packet transmission. It integrates the communication process of the same information and gets rid of irrelevant information, which is able to reduce TCP head size to 50 % and saves the overhead of feedback confirmation. The compressed TCP header length is variable, but it must contain a connection identifier: the one bit for domain and an inspection. It does not use incremental

coding mode. Compression is only implemented in the TCP header, and does not affect the IP header department. Through the use of communication at both ends of the header compression and decompression, SCPS-TP is able to avoid problems caused by the connection status changes, and resolve out of order and return loss of−n behavior by ensuring that each packet is independent decompressed.

3 Simulation Analysis

In order to verify the performance characteristics of SCPS-TP protocol in the space information network, this paper build simulation models based on OPNET software. The process with hierarchical modeling mechanism is divided into three parts: network layer, node layer, SCPS-TP process layer. Network layer is composed by a node and 1.5 G laser communication link, which is used to determine the topology of the network, protocol, application, traffic, and other system settings. Node layer corresponds to the specific space information network equipment, and is abstracted network entity. Satellite nodes and ground nodes model refers to OSI and TCP/IP layered structure, and consist of the plurality of functional modules through the packet stream constituting. Process layer is formed by a finite state machine to achieve SCPS-TP protocol description, the connection between the state machines denote the state transfer [12].

3.1 Simulation Scenario

This paper builds space information simulation platform based on OPNET to verify the performance of the SCPS-TP. The three optional sub-modules SNACK, congestion control and header compression, which constitute the SCPS-TP protocol process model, are established. The design of network model is simplified by a ground transceiver node, a Low Earth Orbit (LEO) satellite node and communication link topology model node. The main function of the satellite node is similar with data forwarding router. Satellite nodes consist of protocol of physical layer, data link layer and network layer. Physical layer includes a plurality of pairs of radio receiver (SAT-RX), a wireless transmitter (SAT-TX). Data link layer uses link layer protocol Advanced Orbiting systems (AOS). Network layer uses the IP protocol. The main function of ground nodes is to generate and receive packets. The three underlying protocol is the same as the underlying protocol of satellite nodes. The transport layer is SCPS-TP protocol. Application layer service is FTP. OPNET provides end-to-end business configuration method. The application on the server is set by configuration object *application config* and the application on the client is set by distribution configuration object *profile config* [13]. The simulation scenario is shown in Fig. 1.

Fig. 1 Simulation scenario

The simulation client node is located in Beijing $(39°.92'N, 116°.46'E)$, the server node is located in Sydney $(33°.55'S, 150°.53'E)$. The design of space-based network refers to Iridium system [14], and makes appropriate simplification and modification. It consists of 48 satellites and uses Walker constellation. Track parameters are generated by the satellite simulation tools package Satellite Tool kit (STK) [15], each satellite runs on each orbit and connects through Intersatellite Link (ISL). Basic parameters are set as shown in Table 1.

In the simulation of three scenarios, the FTP File size is set to 100 M, the round-trip delay is set to 100 ms. In the three scenarios, SNACK module, congestion control module, header compression module is added to SCPS-TP transport layer protocol model separately. Network performance using TCP and SCPS-TP under different Bit Error Rate (BER) conditions is simulated, and transmission throughput performance of the two protocols is compared and analyzed. The simulation parameters are set as shown in Table 2.

3.2 SNACK Simulation Analysis

In the first simulation scenario, the SNACK module is embedded as a child process in the send node and receive node of the transport layer protocol based on simulation platform. The Fig. 2 is the throughput performance change status of the SCPS-TP and standard TCP when using different error rate.

Table 1 Network parameters setting

Orbital altitude (Km)	1,666
Cycle (s)	7,200
Number of satellite	48
Orbit number	6
Orbit inclination	86.2°
Number of inter-satellite links within orbits	2
Number of inter-satellite links between orbits	1

Table 2 Simulation
parameters setting

Parameters	Set
Send cache	32,768 bytes
Receive cache	32,768 bytes
Link bandwidth	2 Mbit/s
Packet size	1,500 bytes
Initialization window	2 MSS
Queue length	128 package
Initialization RTO	1 s
Simulation time	60 s

Fig. 2 SNACK throughput
performance
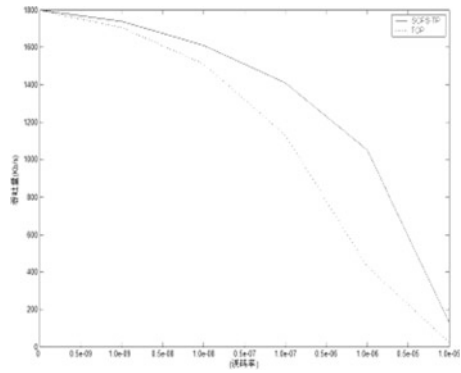
As can be seen from Fig. 2, SCPS-TP and TCP throughput are nearly the same in bit error rate of 0 to $1E − 09$. With the increase of the error rate, SCPS-TP throughput performance changes smaller than TCP. The throughput has significantly reduced only after the error rate is higher than $1E − 06$, which indicates that the SCPS-TP in the data can better maintain its transmission rate than TCP. With the BER increases TCP throughput performance decreased rapidly. This is mainly due to TCP congestion control strategy which makes the loss of all data packets. It takes TCP a cycle at least to detect packet loss, which is slow. As a result, SCPS-TP with SNACK has a better advantage in the large delay and the high bit error rate environment compared with TCP.

3.3 Congestion Control Simulation Analysis

In the second simulation scenarios, SCPS-TP uses the Vegas congestion control algorithm, and TCP uses JV congestion control algorithm. The congestion performance of the two algorithms is analyzed by comparing different bit error rates of the throughput performance.

Fig. 3 Congestion algorithm
throughput performance
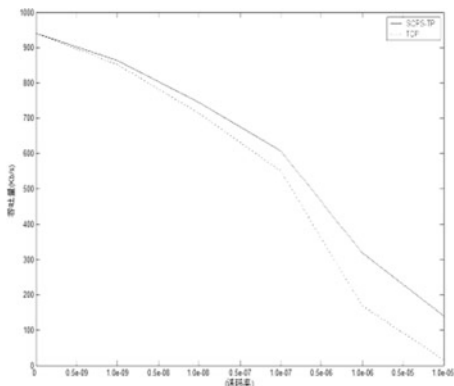
As can be seen from Fig. 3, in the error rate of 0 to $1E - 07$, the throughput of
the two different congestion algorithms reduced slowly and does not change sig-
nificantly. This is mainly because the error rate is low. Packet loss is mainly
caused by link congestion. SCPS-TP throughput decreases more smoothly when
the error rate increases above $1E - 07$ while the TCP throughput declines rapidly.
This is mainly because with the BER increases, packet loss is mainly caused by
link corruption. In this case, the Vegas have better performance than the VJ.

3.4 Head Compression Simulation Analysis

In the third simulation scenario, SCPS-TP uses header compression technology
while TCP does not use any compression technology. Throughput changes under
different error rate as show below:

As can be seen from Fig. 4, the header compression technique in SCPS-TP does
not show much advantage compared to TCP. In different BER, both are sub-
stantially the same throughput. When link is symmetric and bandwidth reduce is
not very low, reducing the size of the data of the response packet does not have a

Fig. 4 Header compression
throughput
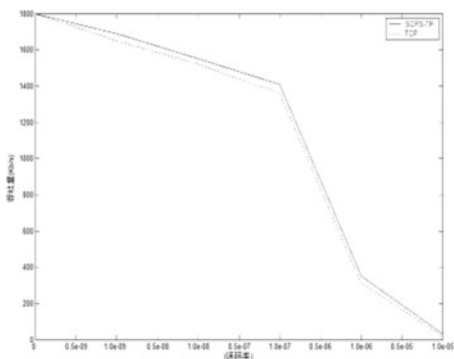
big impact on the throughput. The role of header compression can be reflected, only when the Maximum segment size (MSS) bit error to achieve optimal processing or link has greater asymmetry bandwidth, This requires further research and experiment.

4 Conclusion

This paper establishes the space information network simulation platform. Based on the platform, the research and simulation of SNACK, congestion control algorithms and head compression technology are carried on preliminary. We conclude that: Compared with TCP, SNACK, Vegas congestion control algorithm and header compression of SCPS-TP are more suitable to the large latency, high bit error rate of space information network environment. If the other technologies in SCPS-TP Blue Book are adopted, it would show a better performance than TCP. However, the validity needs further in-depth research, SCPS-TP technology and Transport Layer Security protocol needs improvement, which will be our future work.

References

1. Zhang M, Luo G-C (2008) Reliable transmission control protocol for spatial information networks. J Commun 29(6):63–68
2. Yang L (2011) Research on reliable transport protocol of space communications network. Xidian University, Sian
3. Ding R (2011) Research and simulation of SCPS-TP protocol in space communication. University of Electronic Science and Technology of China, Chengdu
4. CCSDS714 (2006) 0-B-2 Space communications protocol specification (SCPS)-transport protocol (SCPS-TP). Recommendation for space date system standards. Blue Book, Issue 2, CCSDS, Washington, DC
5. Wan P, Wang R-J, Huang W et al (2010) Analysis of the performance of TCP and its extension protocol for space communication. J Spacecraft TT C Technol 29(3):11–16
6. Durst RC, Miller GJ, Travis EJ (1997) TCP extensions for space communication. Wireless Netw Spec Issue: Mobile Comput Networking 5(3):389–403
7. Chen Y, Meng X, Bian C-J (2012) Analysis of protocol and performance of SCPS-TP. Manned Spaceflight 18(1):68–72
8. Fox R (1989) TCP big window and Nak options. RFC 1106
9. Jiang L-Z, Meng X (2008) TCP and SCPS-TP over space networks. In: Proceedings of Second international Symposium on intelligent information technology application. doi:10.1109/IITA:675-678(2008)
10. Brakmo LS, O'Malley SW, Peterson LL (1994) TCP Vegas: new techniques for congestion detection and avoidance. In: Proceedings of SIGCOMM'94: Symposium on communications architectures and protocols
11. Jacobson V (1990) Compressing TCP/IP headers for low-speed serial links. RFC1144
12. Gao S (2010) The large decryption of OPNET modeler simulation and modeling. Publishing House of Electronics Industry, Beijing, pp 15–63

13. Chen M (2004) OPNET network simulation. Publishing House of Tsinghai University, Beijing, pp 188–197
14. Yu J, Zong P (2010) The analysis and simulation of communication network in iridium system based on OPNET. In: Proceedings of information management and engineering (ICIME) 2010 the 2nd IEEE international conference. pp 16–18
15. Yao D-L, Wen D-S, Lin S-H et al (2007) STK and application in simulation of the communication satellite link. Sci Technol Eng 11(7):2700–2702

A Local Elitism Based Membrane Evolutionary Algorithm for Point Pattern Matching

Zhuanlian Ding, Jin Tang, Xingyi Zhang and Bin Luo

Abstract Point pattern matching is a fundamental problem in computer vision and pattern recognition. Membrane computing is an emergent branch of bio-inspired computing, which provides a novel idea to solve computationally hard problems. In this paper, a new point pattern matching algorithm with local elitism strategy is proposed based on membrane computing models. Local elitism strategy is used to keep good correspondences of point pattern matching found during the search, so the matching rate and the convergence speed are improved. Five heuristic mutation rules are introduced to avoid the local optimum. Experiment results on both synthetic data and real world data illustrate that the proposed algorithm is of higher matching rate and better stability.

Keywords Membrane computing · Membrane algorithm · Local elitism · Point pattern matching

Z. Ding · J. Tang (✉) · X. Zhang · B. Luo
School of Computer Science and Technology, Anhui University, Hefei 230601,
Anhui, People's Republic of China
e-mail: ahhftang@gmail.com

Z. Ding
e-mail: dingzhuanlian@163.com

X. Zhang
e-mail: xyzhanghust@gmail.com

B. Luo
e-mail: luobin@ahu.edu.cn

J. Tang · X. Zhang · B. Luo
Key Lab of Industrial Image Processing and Analysis of Anhui Province, Hefei 230039
Anhui, People's Republic of China

1 Introduction

Point pattern matching (PPM) is a fundamental problem in computer vision, and it has been applied widely to image registration, model-based object recognition and stereo vision [1, 2], etc. In general, PPM can be considered as an optimization problem, which was proved to be computationally hard. Many approximate approaches for solving this problem have been proposed, such as spectral methods, graph model, probabilistic relaxation methods and bio-inspired optimization methods [1–3]. Spectral methods are usually not robust to structural corruption and patterns with different sizes. Probabilistic relaxation methods are easy to arrive at local minima. Although bio-inspired optimization methods have to face a large amount of challenges, such as Genetic Algorithm (GA) and Particle Swarm Optimization (PSO), this kind of methods has significant potential to be applied to various optimization problems.

Membrane computing is an emergent branch of bio-inspired computing, with the aim to abstract computing ideas (data structure, operations with data, computing models, etc.) from the structure and the functioning of a single cell or complexes of cells such as tissues and organs. The models obtained are called P systems, which were proved to be a rich framework for handling many problems related to computing. Most variants of P systems are computationally universal [4–6], and can theoretically solve presumably intractable problems in a feasible time (solving NP-complete problems [7, 8]; or even PSPACE-complete problems [9]). Actually, membrane computing provides an approach to solve the intractable problems from a biologically inspired point of view, both in theoretical solutions and in approximate solutions. Membrane computing was proved to be one of the rich sources for inspiring optimization algorithms to solve the intractable problems. The first attempt was done by Nishida [10], where the ingredients of P systems were borrowed, such as membrane structure, transport mechanisms through membranes, etc. Experiment results showed that such new kind of algorithm was more efficient than other classical algorithms on one of the famous NP-complete problems—Traveling Salesman Problem (TSP). Since then, several optimization algorithms based on P systems have been proposed for solving other intractable problems [11–13]. Although a lot of open problems related with membrane computing inspired optimization algorithms need to be investigated, much more attention should be paid to the design of problem-specific algorithms based on P systems, just as suggested by Păun in Ref. [14].

In this paper, a membrane computing inspired optimization algorithm is proposed for solving PPM problem, based on the membrane algorithm proposed by Nishida [10]. The aim of PPM is to find a group of correspondences between two sets, while one correspondence has little effect on another correspondence. So,

it is necessary to conserve good correspondences found during the search. A local elitism strategy is adopted in the proposed algorithm to conserve good correspondences. Experiment results illustrate that such strategy is effective, and the matching rate and the convergence speed are improved. Five new heuristic mutation operators are also introduced to avoid local optimum. Experiment results on both synthetic data and real world data show that the proposed algorithm outperforms its counterpart algorithms, in terms of accuracy, stability and convergence.

The paper is organized as follows. In Sect. 2, the local elitism based membrane evolutionary algorithm for PPM is presented, including a detailed description of the local elitism strategy and five new heuristic mutation operators. Experiment results on synthetic data and real world data are shown in Sect. 3. Conclusions are drawn in Sect. 4.

2 A Local Elitism Based Membrane Evolutionary Algorithm for Point Pattern Matching

PPM can be formulated as follows. Given the model point pattern $P = \{P_i : P_i \in \Re^N, \quad i=1,\ldots m\}$ and the observed point pattern $Q = \{Q_j : Q_j \in \Re^N, \quad j=1,\ldots n\}$, find a map $\hat{f} : P' \to Q'$ ($P' \subseteq P$ and $Q' \subseteq Q$), such that the following matching objective value $Gobj$ is minimized.

$$Gobj = \sum_{(i,j)\in P} \left\{ |d(p_i, p_j) - d(\hat{f}(p_i), \hat{f}(p_j))| + k \times (n_P - n_{P'}) \right\} \qquad (1)$$

where $d(\ldots)$ is the Euclidean distance, $n_{p'}$, n_P are the sizes of P' and P, k is a penalty factor. In this paper, we use the following encoding way: $A = a_1, \ldots, a_{n_p} (0 \leq a_i \leq n_Q)$ encodes a map between P and Q, where $a_i = j$ represents that q_j corresponds to p_i, $a_i = 0$ represents that there is no correspondence between p_i and any point in Q. Note that each a_i should choose different value except 0.

Before describing our algorithm, we first present two key "characteristics" of the algorithm—local elitism strategy and five heuristic mutation operators.

2.1 Local Elitism Strategy

The pseudocode of local elitism strategy is as follows.

Algorithm 1: Local Elitism Strategy

Input: S_1 // the best solution in region 1
Output: M_{set} //M_{set} is used to keep good correspondences obtained in the search
1. **Initialize:** SS, M_{set} //SS is an archive set to store good solutions obtained so far
2. **For** each node i in S_1
3.   **If** satisfies $M_{set}(i)=0$ & $S_1(i) \neq 0$
          $N \leftarrow i$
4.   **End if**
5. **End for**
6. **For** each node i in N
7. $num_i=0$
8.   **For** each node j in N
9.   **If** satisfies $[d(p_i, p_j)/d(\hat{f}(p_i), \hat{f}(p_j))] < \alpha + \tau |M_{set0}|$ // α is a constant and τ is a ratio.
10. $num_i \leftarrow num_i + 1$
11.   **End if**
12.   **End for**
13.   **If** satisfies $num_i \geq \eta |N|$ // η is a ratio
14. $M_{set} \leftarrow i$
15.   **End if**
16. **End for**
17. **If** $rand < \xi$ $(\xi \in (0,1))$
          $M_{set} \leftarrow NULL$
18. **End if**
19. $SS \leftarrow S_1$
20. Output M_{set}

In this strategy, the local elite correspondences are chosen as the ones whose local objective values are small enough. These elite correspondences will be conserved in M_{set} during the search, which will be used to find other good correspondences in further process of search.

2.2 Five Heuristic Mutation Operators

The following five heuristic mutation operators are introduced to avoid local optimum.

Exchange mutation-1. Perform two points exchange operation with probability P_{e1}. The two points chosen have the highest mutation probability and the second highest mutation probability, respectively.

Exchange mutation-2. Execute two points exchange operation with probability P_{e2}. One point chosen has the highest mutation probability, while another point is chosen randomly.

Exchange mutation-3. Select two points randomly to exchange each other with probability P_{e3} to avoid local optimum.

Replace mutation. Select the highest mutation probability point and replace it with one picked from those unsigned points with probability P_r.

Zero-mutation. Select the highest mutation probability point and set its value as 0 with probability P_m.

The mutation probability in the above five operators is calculated as follows. Given an A with n_P points, we first calculate the matching objective value of each point as the following:

$$Pobj(i) = \sum_{j \in P_{set}} \left| d(p_i, p_j) - d(\hat{f}(p_i), \hat{f}(p_j)) \right| \Big/ |P_{set}| \qquad (2)$$

where $|P_{set}|$ is the size of the set which contains all nonzero points of A. Note that $Pobj(i) = k$ if $\hat{f}(p_i) = 0$ satisfies, where k is the penalty factor described above. Mutation probability of each point is calculated as follows:

$$P(i) = Pobj(i) \Big/ \sum_{j=1}^{n_P} Pobj(j) \qquad (3)$$

With these operators, it is heuristic to select mutation points to mutate under a certain probability.

2.3 Local Elitism Based Membrane Evolutionary Algorithm for Point Pattern Matching

Based on the local elitism strategy and the heuristic mutation operators described above, the proposed algorithm can be summarized as follows.

Algorithm 2: Local Elitism Based Membrane Evolutionary Algorithm for PPM

Input: membrane number x, M_{set} and temperature T_i of region i and initial solutions.
Output: S_{best}
1. Initialize: SS, M_{set} //SS is an archive set and M_{set} keeps the better correspondences
2. **While** iteration
3.   **For** each region i
4. $s_i \rightarrow s_i'$ //Execute evolution rules, s_i is the initial solution in region i
                //Note that the points in M_{set} are kept during the evolution
5.   Calculate $\Delta \leftarrow Gobj(s_i') - Gobj(s_i)$
6.   If $\Delta < 0$, $s_i \rightarrow s_i'$ else $s_i \rightarrow s_i'$ with the probability $\exp(-\Delta/T_i)$
7.   **End for**
8.   **For** each region i
9.   //Execute communication rule
    Send the best solution in region i to region i-1 and the worst solution to region i+1
10.   **End for**
11.   Select the best solution in each region as the next generation
12.   Execute local elitism strategy to update M_{set}
13. $S_{best} \leftarrow$ the best solution in archive set SS
14.   Iteration $\leftarrow$ iteration+1
15. **End while**

Fig. 1 The membrane
structure of the proposed
algorithm
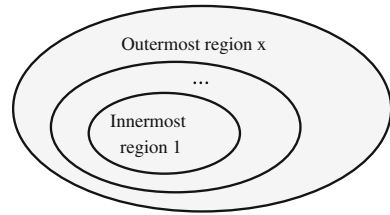
In view of the characteristics of PPM problem, the above algorithm is designed
for solving such problem, based on the membrane algorithm proposed by Nishida.
Figure 1 shows the membrane structure of the algorithm.

3 Experimental Results

3.1 Experiment Data

The synthetic data are created as follows: a set P consisting of 40 points is
generated randomly in 512×512 regions; Two kinds of observed point pattern
Q are generated by adding Gaussian noise with different variance σ to P, and
adding outlier noises with different ratio r to P, respectively. In the second case of
generating Q, we choose $n_Q = (1+r) \times n_P$. The real world data we use are YORK
and MOVI image data. On YORK image data, the second frame is matched to
some subsequent frames (from the fourth to the twelfth), where 42 landmark points
are marked on each image. On MOVI image data, the first frame is used as model
image and some subsequent frames (from the fourth to the eighth) as observed
images, and we extract 120 Harris corner points on each image.

The parameters in our experiments are used as follows: $T_i = T_x \times q^{x-i}$, where the
cooling coefficient $q = 0.8$ and the temperature of outermost region $T_x = 100$. We set
$x = 30$, $\eta = 0.9$, $\alpha = 0.08$, $\tau = 0.005$, $\xi = 0.02$, $P_{e1}=1$, $P_{e2}=0.4$, $P_{e3}=0.08$, $P_r=0.5$,
$P_m=0.02$. The maximum iteration number is 200 and the number of trials is 20.

3.2 Accuracy

We compare our algorithm with General Membrane Algorithm (GMA), we call
the membrane algorithm proposed by Nishida [10] GMA for distinguishing it with
our algorithm), Particle Swarm Optimization Based Point Pattern Matching
(PSOPPM) [3], Reweighted Random Walks for Graph Matching (RRWM) [1] and
Weighted Voting (WV) [2]. Fig. 2a shows the average matching rate (also called
accuracy) over 20 trials with different Gaussian noise and Fig. 2b presents the
accuracy with different outliers. The matching rate is measured by the number of
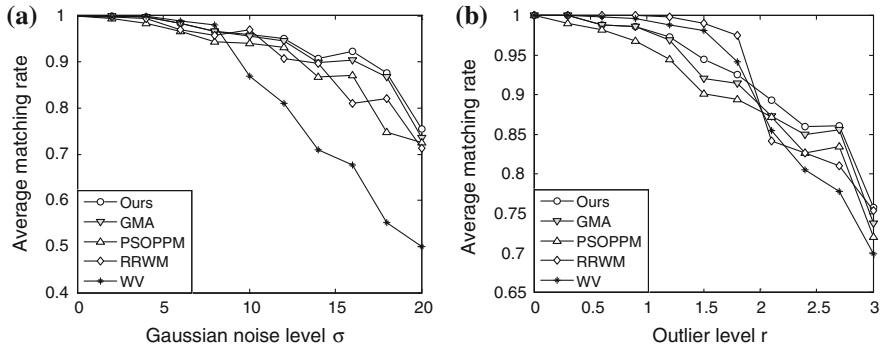
Fig. 2 Comparison of matching rate on synthetic data

detected true correspondences divided by the total number of ground truths. From Fig. 2, it can be found that: (1) our algorithm outperforms the other algorithms in the matching rate under Gaussian noise and outlier noise. Compared with GMA, the proposed local elitism strategy in our algorithm significantly improves the matching rate. (2) RRWM and WV have a similar performance with our algorithm under small noise; however, our algorithm outperforms these algorithms when there exists larger noise.

Our algorithm also achieves a higher matching rate on real world data. Fig. 3a, b illustrate the results on YORK images and MOVI images, respectively.

3.3 Stability and Convergence

In order to evaluate the stability of our algorithm under different noises, the standard deviation (*std*) of matching rate on synthetic data are calculated. Fig. 4a, b show the experiment results under different Gaussian noise and different
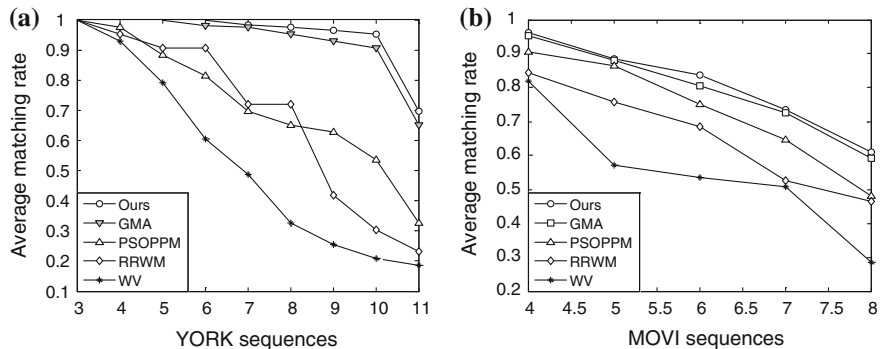
Fig. 3 Comparison of matching rate on YORK and MOVI image data
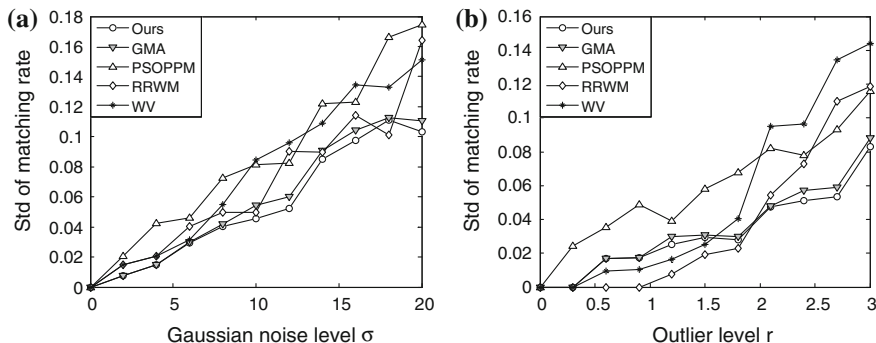
Fig. 4 Comparison of statistics *std* on synthetic data

Table 1 Comparison of statistics *std* on real world image data

Images	Points	Method	Accurancy ($\pm$ *std*)
YORK(2–7)	42	*Ours*	*98.32 % $\pm$ 0.0341*
		GMA	97.67 % $\pm$ 0.0398
		PSOPPM	70.11 % $\pm$ 0.1661
YORK(2–11)	42	*Ours*	*69.79 % $\pm$ 0.0477*
		GMA	65.12 % $\pm$ 0.0636
		PSOPPM	32.68 % $\pm$ 0.1988
MOVI (1–6)	120	*Ours*	*83.71 % $\pm$ 0.0427*
		GMA	80.65 % $\pm$ 0.0595
		PSOPPM	75.03 % $\pm$ 0.1775
MOVI (1–8)	120	*Ours*	*61.17 % $\pm$ 0.0509*
		GMA	59.22 % $\pm$ 0.0851
		PSOPPM	48.20 % $\pm$ 0.1991

outlier ratio over 20 trials, respectively. From these Figures, it can be seen that the proposed algorithm outperforms the other algorithms in stability. More specifically, the membrane computing inspired methods (including GMA and the proposed algorithm) are obviously more robust to outliers than PSOPPM, RRWM and WV, and the proposed algorithm which incorporates the local elitism strategy is consistently superior to the old membrane algorithm GMA. In particular, the other algorithms are comparable to our algorithm under less Gaussian noise level as shown in Fig. 4a, however, with increasing noise, the stability of the other approaches decreases considerably. Fig. 4b also shows a similar result.

As shown in Table 1, the results obtained by the proposed algorithm are more stable (with smaller *std*) than those obtained by GMA and PSOPPM.
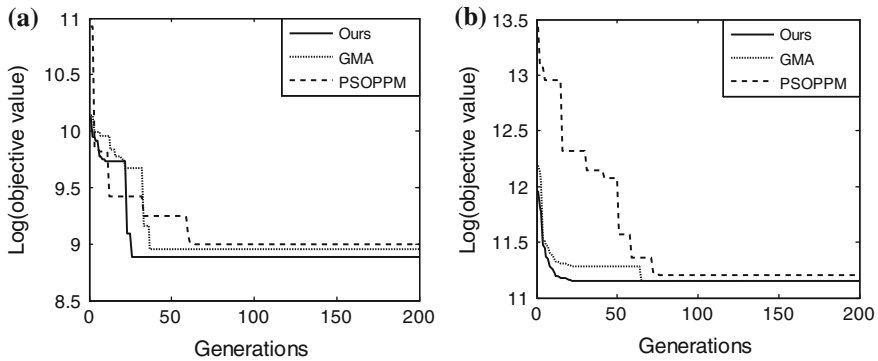
Fig. 5 Comparison of convergence speed

Fig. 5a, b show the convergence curves of our algorithm on synthetic data and MOVI image data, respectively. Obviously, the proposed algorithm achieves faster convergence than GMA and PSOPPM. This may be caused by the local elitism strategy proposed in our algorithm.

4 Conclusion

In this paper, based on the membrane algorithm proposed by Nishida, we present a local elitism based membrane evolutionary algorithm for PPM problem. The local elitism strategy is proposed for conserving good correspondences found during the search, which is proved to be an effective strategy for improving the matching rate and the convergence speed. Experiment results done on synthetic data and real world data, illustrate that the proposed algorithm can obtain a higher matching rate and a better stability for PPM problem. As shown in Fig. 1, the membrane structure of the proposed algorithm consists of several regions and each region can work in parallel. Therefore, it is expected to find a parallel hardware suitable for implementing such algorithm. Graphic Processing Unit (GPU) has significant potential to provide the right parallel hardware.

References

1. Cho M, Lee J, Lee KM (2010) Reweighted random walks for graph matching. European conference on computer vision. Heraklion, Greece, pp 492–450
2. Yuan Y, Pang YW, Wang KQ, Shang MY (2012) Efficient image matching using weighted voting. Pattern Recogn Lett 33(4):471–475
3. Yin PY (2006) Particle swarm optimization for point pattern matching. J Vis Commun Image Represent 17(1):143–162
4. Pan LQ, Păun G (2009) Spiking neural P systems with antispikes. Int J Comput Commun Control IV(3):273–282
5. Pan LQ, Păun G (2010) Spiking neural P systems: an improved normal form. Theoret Comput Sci 411(6):906–918
6. Wang J, Hoogeboom HJ, Pan LQ, Păun G, Pérez-Jiménez MJ (2010) Spiking neural P systems with weights. Neural Comput 22(10):2615–2646
7. Pan LQ, Martín-Vide C (2005) Solving multidimensional 0–1 knapsack problem by P systems with input and active membranes. J Parallel Distrib Comput 65(12):1578–1584
8. Pan LQ, Martín-Vide C (2006) Further remark on P systems with active membranes and two polarizations. J Parallel Distrib Comput 66(6):867–872
9. Alhazov A, Martín-Vide C, Pan LQ (2003) Solving a PSPACE-complete problem by recognizing P systems with restricted active membranes. Fundamenta Informaticae 58(2):67–77
10. Nishida TY (2004) An application of P systems: a new algorithm for NP-complete optimization problems. In: Callaos N et al (eds) Proceedings of the 8th world multi-conference on systems, cybernetics and informatics, vol 5, pp 109–112
11. Zhang GX, Pan LQ (2010) A survey of membrane computing as a new branch of natural computing. Chin J Comput 32(2):208–214
12. Huang L, Suh IH, Abraham A (2011) Dynamic multi-objective optimization based on membrane computing for control of time-varying unstable plants. Inf Sci 181:2370–2391
13. Xiao JH, Zhang XY, Xu J (2012) A membrane evolutionary algorithm for DNA sequence design in DNA computing. Chin Sci Bull 57(6):698–706
14. Păun G (2007) Tracing some open problems in membrane computing. Romanian J Inf Sci Technol 10(4):303–314

Prediction of Coal Calorific Value Based on a Hybrid Linear Regression and Support Vector Machine Model

Kelei Sun, Rongrong Gu and Huaping Zhou

Abstract The gross calorific value (GCV) is an important property defining the efficiency of coal. There exist a number of correlations for estimating the GCV of a coal sample based upon its proximate and ultimate analyses. These correlations are mainly linear in character although there are indications that the relationship between the GCV and a few constituents of the proximate and ultimate analyses could be nonlinear, which has made artificial intelligence models as a useful tool for a more accurate GCV prediction. This paper focuses on an innovative method of GCV prediction using combination of Multivariate Linear Regression (MLR) as predictor and Support Vector Machine (SVM) as an error correction tool based on proximate and ultimate analyses. The GCV have been predicted using the MLR, ANN and the hybrid MLR–SVM models. In the analysis root mean squared error have been employed to compare performances of the models. Results demonstrated that both models have good prediction ability; however the hybrid MLR–SVM has better accuracy.

Keywords Gross calorific value · Error correction · SVM · Regression analysis

K. Sun (✉) · R. Gu · H. Zhou
School of computer Science and Engineering, Anhui University of Science and Technology, Huainan 232001, China
e-mail: klsun@aust.edu.cn

R. Gu
e-mail: rrgu@aust.edu.cn

H. Zhou
e-mail: hpzhou@aust.edu.cn

1 Introduction

The calorific value is one of the important indexes to evaluate the coal quality, calculating the heat balance, coal consumption and thermal efficiency in the combustion process. Calorific value is a rank parameter, but is also dependant on the mineral composition. The experimental determination of gross calorific value (GCV) of coal is a cost intensive process, as it requires special instrumentation and highly trained analyst to operate it. For this reason, a number of empirical formulas have been developed for the prediction of GCV based on proximate and ultimate analyses. Given et al. [1] used theoretical physical constants to develop an equation to calculate GCV from elemental composition; Mason and Gandhi [2] used regression analysis and data from 775 coals to develop an empirical equation that estimates GCV of coal from C, H, S and ash [1, 2].

The essence of GCV is a concentrated expression of moisture, volatile matter, ash, fixed carbon and other elements. The moisture evaporates in coal combustion, which consumes calories. The ash can affect GCV, which reduces the proportion of combustible portion. There are nonlinear relationship between GCV and mineral composition of coal [3–5]. It is inappropriate to characterize this non-linear relationship in a linear model. Therefore, it is important to study the nonlinear predicted method of GCV.

2 Experimental Data and Correlation Analysis

Data used to test the proposed approaches are from Huainan Mining Industry Coal Quality database. The samples with high ash as well as the samples with same proximate and ultimate analysis were excluded from the database. A total of 52 set of coal sample analysis were used. The proximate analysis data of coal with experimental GCV are shown in Table 1.

It is well-known that ash, moisture, volatile matter and fixed carbon collectively decide the nature of coal. It was observed that ash and moisture have negative effects on GCV. Research results show that there is a good correlation between GCV and coal ingredient, such as moisture and ash [6, 7]. The paper regards moisture and ash as two independent variables, GCV as a dependent variable affected by the former two parameters.

2.1 Effects of Moisture on GCV

Moisture evaporates in the combustion process, which takes part of the heat away. So the GCV is decreased as the moisture content increases, that is to say there is an inverse correlation between GCV and moisture content. To show the impacts of

Table 1 Proximate analysis data of coal with experimental GCV

No.	A_d (%)	M_t (%)	$Q_{net.ad}$ (MJ·kg^{-1})	No.	A_d (%)	M_t (%)	$Q_{net.ad}$ (MJ·kg^{-1})
1	26.97	7.3	21.38	27	26.44	5.6	22.79
2	26.63	6.7	22.29	28	25.12	6.9	22.60
3	25.61	6.3	22.50	29	23.70	6.4	23.72
4	25.61	6.3	22.50	30	24.50	5.1	23.57
5	26.39	7.5	21.76	31	23.41	5.6	22.60
6	26.58	6.9	21.81	32	24.22	5.6	23.38
7	30.61	5.0	21.16	33	28.16	5.1	22.00
8	29.07	5.2	21.21	34	25.30	6.4	22.70
9	25.58	5.0	23.02	35	29.51	5.6	21.16
10	27.40	5.8	21.93	36	27.61	5.1	21.89
11	24.18	5.8	23.15	37	27.34	5.1	22.04
12	26.00	5.4	22.55	38	23.48	7.9	22.60
13	26.93	5.7	21.90	38	28.73	5.9	21.71
14	26.52	6.1	22.25	40	24.41	5.5	22.16
15	25.56	5.2	23.56	41	30.61	5.0	21.16
16	25.99	5.4	22.43	42	27.83	5.8	21.67
17	26.97	7.3	21.38	43	24.23	3.7	23.51
18	24.24	4.4	23.44	44	30.61	5.0	21.16
19	29.07	5.2	21.21	45	23.09	5.3	24.30
20	24.41	5.5	22.16	46	23.45	5.2	23.98
21	28.38	5.5	21.73	47	30.55	4.7	20.95
22	26.58	5.7	22.22	48	28.67	5.0	21.44
23	28.75	5.3	21.84	49	25.12	4.8	23.82
24	26.56	5.4	22.43	50	24.06	5.8	23.17
25	30.06	5.3	21.22	51	25.30	4.6	23.41
26	28.24	4.0	22.53	52	23.43	5.7	24.27
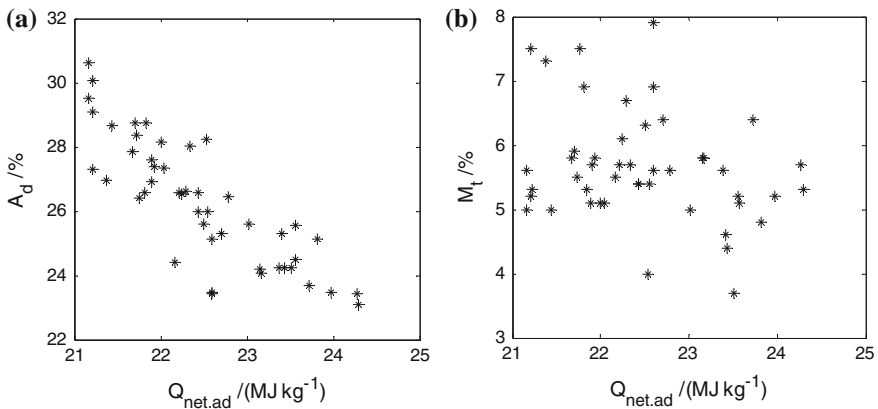
Fig. 1 Effects of ash and moisture contents on GCV of coal, **a** ash contents **b** moisture contents

ash contents of coals on GCV, the GCV are plotted as a function of the added value of ash, which are shown in Fig. 1a.

2.2 Effects of Ash on GCV

During the combustion process the vast majority of mineral substance decomposes with absorption of heat. The more the mineral content is, the more heat is absorbed in pyrolysis. While the ash content is increased by 1 %, the GCV would fall by 300–380 J/g. The relationships between GCV and ash content are shown in Fig. 1b.

3 Traditional Predicted Methods

3.1 Multivariate Regression Analysis

There are two different methods for regression analysis: (1) establish a regression equation by the least squares method; (2) rapid regression analysis taking advantage of the EXCEL tool. Essentially, the two methods are the same. The training sample set $\{X_t\ Y_t\}$ is composed of raw data from the record. Input vector X_t is the two-dimensional vector, which represents ash content and moisture content of each sample. Output vector Y_t is the experimental GCV of coal samples. A MLR predicted model is constructed. Based on training sample set $\{X_t, Y_t\}$, the regression parameter is calculated by fitting algorithm.

In this paper, MLR is run on MATLAB software platform. Input the training sample set, then a MLR model is

$$Q_{net\ ar} = -0.3529A_d + 0.3173M_{ad} + 33.4628 \tag{1}$$

In fact, there is not strict linear relationship between GCV and ash/moisture. So it is difficult to obtain the high accuracy if only the linear regression model is used.

3.2 Artificial Neural Network

There are stronger fault tolerance and non-linear processing capability in artificial neural network. Back-Propagation (BP) network is an artificial neural network model which is used most widely. It adopts supervised learning algorithm. The network learning process includes two stages, feed-forward calculation and reverse weight adjustment. In essence, the learning rule for BP network adopts gradient descent. Its weight doesn't changes along negative gradient direction of error function until error meets the requirement of output. In fact, the BP network training process is a process to find the stable weight values and the best network architecture by adjusting the weights of each input value.
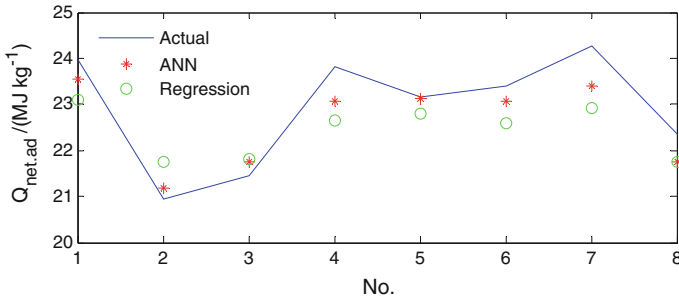
Fig. 2 Experimental GCV and predicted GCV

There are former 44 learning samples for training, which is shown in Table 1. The three-layer BP network is adopted. The input layer has two nodes, which come from moisture and ash content data of industrial analysis data. There are six neurons selected in the hidden layer, which respects logsig function as transformation function. The output layer only contains a node, the GCV, which chooses the purelin function as transformation function. Other parameters are: training maximum: 5,000, maximal error of network training objective: 1.0×10^{-3}, momentum factor: 0.7 and the default value is used for other parameters.

Another 8 group of measured data in Table 1 are tested for predicted samples by the two before-mentioned models. The predicted results are consistent with the measured ones, which are shown in Fig. 2.

4 Hybrid LR-SVM Model

In order to improve the prediction accuracy, the SVM algorithm is used for amending the prediction error of the former two models. The basic thought of SVM regression algorithm is described as follows. Given a training set of l data samples, and each sample is a n-dimensional input vector. $\varphi(\cdot)$ is a nonlinear function mapping the input space into a higher dimensional space, then the optimal linear regression function is constructed in this space [8–10].

$$f(x) = \varpi^T \varphi(x) + b \tag{2}$$

LSSVM is an extended version of standard SVM in which solutions can be solved by equality constraints instead of inequality constraints. Thereby the optimization problem can be described as

$$\begin{cases} \min J(w, e) = \frac{1}{2} w^T w + \frac{\xi}{2} \sum_{i=1}^{l} e_i^2 \\ s.t. y_i = w^T \varphi(x_i) + b + e_i \end{cases} \tag{3}$$

where w is a coefficient vector which can indicate function complexity; b is a constant term; e is a residual vector; ξ is a regulation constant which can adjust the complexity of model and the training error. ξ can take an intermediate value so that the income function has better generalization ability. The value of ξ is in inverse proportion to the regression error of model. Then Lagrange multipliers are added and the Kuhn–Tucker conditions are listed which can be written as the following set of linear equations:

$$\begin{bmatrix} 0 & Y^T \\ Y & \Omega + \gamma^{-1}I_l \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{4}$$

where $a = [a_1 \ldots a_n]^T$; $y = [y_1 \ldots y_n]^T$; $Y = [1 \ldots 1]^T$; Ω is square matrix in which element satisfies Mercer condition; γ is an adjustment constant.

According to Mercer condition, the optimal parameters a, b can be calculated by the least squares method. Finally LSSVM regression model is obtained as:

$$y(x) = \sum_{i=1}^{l} a_i K(x, x_i) + b \tag{5}$$

A minimal subset of the sum of empirical risk and fiducial range is adopted by LSSVM, which based on structural risk minimization (SRM) formula. A 2-norm of ξ is selected as the loss function in regression algorithm. The quadratic programming problem is replaced by the solution of a system of linear equations in this algorithm, so that the complexity of algorithm can be greatly reduced, and the interference rejection of modeling process can be improved. Especially, there is strong generalization ability shown in small-sample learning problems, which can effectively solve the over study problem in the error correction process, so that the accuracy of amendment can be guaranteed.

The concrete steps to amend the prediction error by LSSVM algorithm are as follows:

(1) Input industrial analysis data of the sample, and establish regression analysis prediction model, then a calorific predicted value Q_Y of the sample can be obtained;
(2) Compare predicted value Q_Y with experimental value Q, then the prediction error ΔQ_Y of regression analysis can be calculated;
(3) Analyze the prediction error ΔQ_Y by LSSVM algorithm again, and establish a regression analysis pre-compensation model to amend the former results. The prediction process is shown in Fig. 3.

In this paper, the hybrid LR-SVM algorithm is achieved by the matrix functions in Matlab. The parameters are selected as follows: regulation constant γ: 500; kernel function: Radial Basis Function (RBF) which has high regression accuracy; Radial Basis Parameter δ: 0.1. Input the training sample set and the testing sample set into error correction sub-model, the results is shown in Table 2.

Fig. 3 Linear regression model with LSSVM error correction
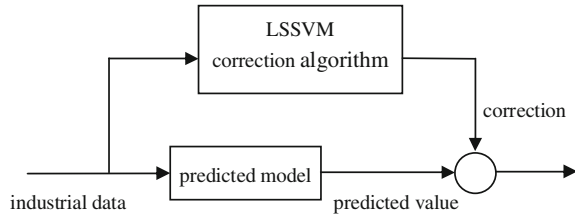
Results of the three models have been compared and discussed statistically and their respective computation times are illustrated in Tables 2 and 3. All the experiments carried out in this endeavor, have been run in an Intel 2.40 GHz environment with 4 GB RAM memory. The results show that, only overall estimate to GCV can be made in multivariate linear regression models. ANN model has nonlinear learning ability, and its prediction accuracy is better than linear regression model. It can be seen that the hybrid LR-SVM model has higher coefficient of determination and fewer errors than the MLR model and ANN model. Thus, the hybrid LR-SVM model produced more accurate results.

As seen in Table 3, the RMSE and MRE have decreased by using SVM error correction method. This means that SVM error correction model is able to compensate error of predicted model, hence, the positive effect of the SVM error correction model is more sensible. In other words, SVM error correction method can increase model prediction efficiency.

Table 2 Experimental GCV and predicted GCV

No.	Observed value	MLR		ANN		LR-SVM	
		Predicted value	Error	Predicted value	Error	Predicted value	Error
1	23.98	23.11	0.87	23.54	0.44	23.84	0.14
2	20.95	21.77	−0.82	21.19	−0.24	21.15	−0.20
3	21.44	21.81	−0.37	21.76	−0.32	21.69	−0.25
4	23.82	22.66	1.16	23.08	0.74	23.74	0.08
5	23.17	22.79	0.38	23.13	0.04	22.96	0.21
6	23.41	22.58	0.83	23.08	0.33	23.33	0.08
7	24.27	22.93	1.34	23.39	0.88	24.10	0.17
8	22.34	21.76	0.58	21.76	0.58	22.22	0.12

Table 3 Comparison of three models by statistics estimators

Estimator	MLR	ANN	LR-SVM
MRE	0.0346	0.0195	0.0068
RMSE	0.9163	0.5506	0.1782
R^2	0.9141	0.9587	0.9963

5 Conclusions

The relationship between GCV and various components of coal is very complex, so traditional MLR method is difficult to achieve the expected effect. Meanwhile, because the number of the coal proximate analysis data is limited, the prediction accuracy of the existing ANN predicted model is lower. In this paper, the multivariate linear regression method is used to reduce the prediction error of the system, and the LSSVM regression algorithm is adopted to improve the complex nonlinear relationship handling ability. Statistical error estimators have been used to compare performance of the MLR model, the ANN model and the hybrid LR-SVM model. Errors estimated by the RMSE and MRE decreased as R^2 increased by implementing the SVM error correction method. Generally, the hybrid LR-SVM model has good ability to predict GCV.

References

1. Given PH, Weldon D, Zoeller JH (1986) Calculation of calorific values of coals from ultimate analyses: theoretical basis and geochemical implications. Fuel 65:849–854
2. Mason DM, Gandhi KN (1983) Formulas for calculating the calorific value of coal and chars. Fuel Process Technol 7:11–22
3. Majumder AK et al (2008) Development of a new proximate analysis based correlation to predict calorific value of coal. Fuel 13:3077–3081
4. Patel SU et al (2007) Estimation of gross calorific value of coals using artificial neural networks. Fuel 3:334–344
5. Mesroghli S, Jorjani E, Chehreh Chelgani S (2009) Estimation of gross calorific value based on coal analysis using regression and artificial neural networks. Int J Coal Geol 1:49–54
6. Maixi Lu, Zhou C (2009) Coal calorific value prediction with linear regression and artificial neural network. Coal Sci Technol 37:117–120
7. Jiang W, Hongqi W, Qu T (2011) Prediction of the calorific value for coal based on the SVM with parameters optimized by genetic algorithm. Thermal Power Gener 40:14–19
8. Dai L-K, Yao X-G (2004) A least squares SVM algorithm for NIR gasoline octane number prediction. Intelligent control and automation, vol 4. WCICA, pp 3779–3782
9. Balabin RM, Lomakina EI (2011) Support vector machine regression—an alternative to neural networks (ANN) for analytical chemistry. Comparison of nonlinear methods on near infrared (NIR) spectroscopy data. Analyst 136(8):1703–1712
10. Balabin RM, Safieva RZ, Lomakina EI (2007) Comparison of linear and nonlinear calibration models based on near infrared (NIR) spectroscopy data for gasoline properties prediction. Chemometr Intell Lab Syst 88(2):183–188

A New Attempt for Satisfiability Problem: 3D DNA Self-Assembly to Solve SAT Problem

Xuncai Zhang, Ruili Fan, Yanfeng Wang and Guangzhao Cui

Abstract The computational speed of an algorithm is very important to NP-hard problems. The 3D DNA self-assembly algorithm is faster than 2D, while 2D is faster than traditional algorithms because DNA molecule owns high parallelism and density. In this paper we mainly introduced how the 3D DNA self-assembly solves the SAT problems. Firstly, we introduced a non-deterministic algorithm. Secondly, we designed seed configuration and different types of DNA tiles which are needed in the computation. Lastly, we demonstrated how the 3D DNA self-assembly solves the SAT problem. In this paper, 3D DNA self-assembly algorithm has a constant tile types, and whose computation time is linear.

Keywords 3D DNA self-assembly · Satisfiability problem · DNA computing

1 Introduction

In many NP-complete problems, SAT problem (satisfiability problem) is called the seed of the other problems [1], which is applied in many fields such as hardware test, artificial intelligence and very important in the research of circuit design, FPGA routing and so on. So the research of the SAT problem has a lot of theoretical and practical values.

NP-complete problems are integral to our everyday lives, but there were no one efficient algorithms to solve it. Adleman [2] succeeded in solving an instance of the Hamilton path problem in 1994, which indicated it was feasible to use DNA molecular and solve NP-hard problems by Biological nanometer technology [3] in polynomial time. Many researchers have done lots of work in the researches of

X. Zhang (✉) · R. Fan · Y. Wang · G. Cui
College of Electrical and Electronic Engineering,
Zhengzhou University of Light Industry, Zhengzhou 450002, China
e-mail: zhangxuncai@163.com

891

SAT problem by using DNA molecules. For example: Lipton solved connected network and regular circuit problems [4] in 1995; in 2000 Sakamoto et al. [5] ingeniously used DNA molecular computation methods of hairpin form to solve CNF-sat problem; Braich et al. [6] gave a half automation assembly of DNA computing model to solve SAT problem which have 20 variables in 2002.

"What is a minimal tileset that can accomplish this goal?" and "what is the minimum assembly time for this system?" are two important questions about creating shapes or computing functions in self-assembling systems. Adleman has emphasized studying the number of steps which takes for an assembly to complete and the minimal number of tiles necessary to assemble a shape [7, 8]. Brun had extended these questions to apply to system [9–11]. In this paper we did some researches to reduce the complexity around these two questions.

2 Related Work for 3D DNA Self-Assmebly

In this section, we focused on the preparation work of solving the SAT problem by 3D DNA self-assembly. Firstly, we imported a non–deterministic algorithm in the SAT problem; secondly, we introduced the formula table which helps us to design the seed configure and leads to the seed configuration; lastly, the design of the 3D tiles was explained.

2.1 Non-Deterministic Algorithm

A non-deterministic algorithm implies there are some non-deterministic choices at some steps of an algorithm. The non-deterministic algorithm of SAT problem was given as below. Noticed that step 2 was the non-deterministic step.

Non-Deterministic Algorithm (X, C):

(1) For each $x_i \in X(n)$ {
(2) Assign $x_i : f(i) \rightarrow \{0, 1\}$
(3) Check all c_i, $c_i \in C(n)$,   if $c_i = 0$
(4) Break and return failure   }
(5) If all $x_i \in X(n)$ are assigned
(6) Return success and output x_i
(7) Else return failure

2.2 Formula Table and Seed Configuration

Any SAT problems can be expressed by Boolean SAT formulas, and any keys of SAT problem is equal to the answers of Boolean SAT formula, so we put attentions to seek the answer of Boolean SAT formula.

A Boolean SAT formula is formed by variables $x_1, x_2, \ldots\ldots, x_m$ and some clauses $c_1, c_2, \ldots\ldots, c_n$. Each clause is independent and comprises of different letters. A SAT formula is a conjunction of clauses and equals 1 when there are some assignments of variables to make all clauses satisfied.

Build a Formula Table. In the SAT formula, we found out each variable have three situations in a clause, so we gave each variable three values: -1, 0 and 1 which represent the variable's solutions in a clause. "-1" represents the reverse side, "0" represents no mentioned and "1" represents the obverse side. Then we can design a table format that could promote a transition to DNA assembly. For example a Boolean SAT formal with 5 variables and 8 clauses was given as below:

$$(x_1 + x_2 + \neg x_3)(\neg x_2 + x_4)(\neg x_1 + \neg x_5)(\neg x_1 + x_2 + x_3 + \neg x_4 + x_5)(\neg x_3)$$
$$(x_2 + x_5)(\neg x_5)\,(x_1)$$

Then the table format of the Boolean SAT formula was just like in Table 1.

The Seed Configuration. The processes of self-assembling start from the seed configuration, which is a beforehand DNA structure constructed artificially, carrying the information of a problem. Once the seed configuration is built, the self-assembly process starts. Before that there was a fact worth to be noticed, in our design, the titles were assumed that they cannot be rotated.

In our design, the seed configuration was made up by some 3D adjacent tiles. Through appropriate placement of every 3D tiles, we can get the values of every variable in each clause. In order to have a stereo feeling of the assembly progresses clearly, we captured a plane, which was the combination of 3D DNA tiles' σ_z surfaces. Figure 1 showed an example of seed configuration for the 3D self-assembly model.

2.3 Design of 3D DNA Tiles

Winfree said if 3D structures are allowed, self-assembly is still universal. In 2010 Minqi Lin [12] invented a 3D DNA self-assembly model to solve the Graph Vertex Coloring problem. We solved the SAT problem with the help of his inspiration.

Table 1 Formula table, an example of the Boolean SAT formula

x_5	x_4	x_3	x_2	x_1	
0	0	−1	1	1	C_1
0	1	0	−1	0	C_2
−1	0	0	0	−1	C_3
1	−1	1	1	−1	C_4
0	0	−1	0	0	C_5
1	0	0	1	0	C_6
−1	0	0	0	0	C_7
0	0	0	0	1	C_8
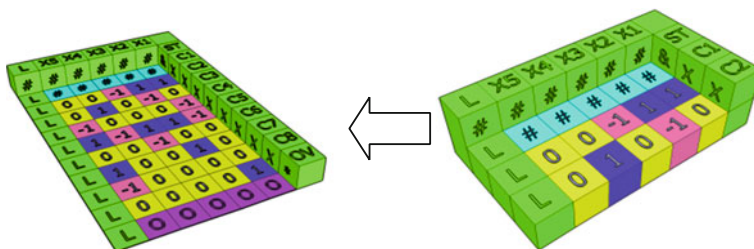
Fig. 1 A example of seed configuration for 3D self-assembly mode, which Boolean SAT formal is $Y = (x_1 + x_2 + \neg x_3)(\neg x_2 + x_4)(\neg x_1 + \neg x_5)(\neg x_1 + x_2 + x_3 + \neg x_4 + x_5)(\neg x_3)(x_2 + x_5)$ $(\neg x_5)(x_1)$

3D DNA tiles are built by DNA strands with unpaired six sticky ends, which are attached with other tiles having the Watson–Crick complementary sticky ends. We showed the 3D DNA molecular model in Fig. 2a and a hexahedron in Fig. 2b which is an abstraction of 3D DNA molecular model. Six labels of the cube stand for a particular kind of sticky ends on each surface respectively. From Fig. 2, we knew that a 3D DNA tile can be denoted by a 6-tuple $\left(\sigma_x, \ \sigma_{-x}, \ \sigma_y, \ \sigma_{-y}, \ \sigma_z, \ \sigma_{-z} \right) \in \sum^6$, which means the six variables can decide a tile. In this paper, two sticky ends can match and legate when they are the identical labels. We designed five types' tiles—the boundary tile, input and output tile, check tile and pass tile in this passage, which helped us to complete the assembly duty.

Input tiles: There are two values for each input variable, so we designed the input tiles as: $\left(\sigma_x = \sigma_{-x} = \&; \ \sigma_y = \{0, 1\}; \ \sigma_{-y} = \#; \ \sigma_z = \text{null}; \ \sigma_{-z} = \# \right)$ (Fig. 3a showed the details).

Output tiles: Output is corresponding to the input, so there must be two types of output tiles, then we designed the output tiles like this: $\left(\sigma_x = \sigma_{-x} = *; \ \sigma_y = \sigma_{-y} = \{0, 1\}; \ \sigma_z = \text{null}; \ \sigma_{-z} = 0 \right)$ (Fig. 3b showed them). The output tiles' duty is to put out the information of input tiles. So the value of σ_y surface is passing the σ_{-y} surface's value.

Pass tiles: Pass tiles' duty are just passing the values, so σ_x surface passes σ_{-x} the surface's label and σ_y surface passes σ_{-y} the surface's label. We designed six

Fig. 2 a Molecular model of 3D tile. **b** Hexahedron model
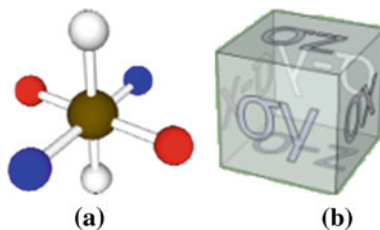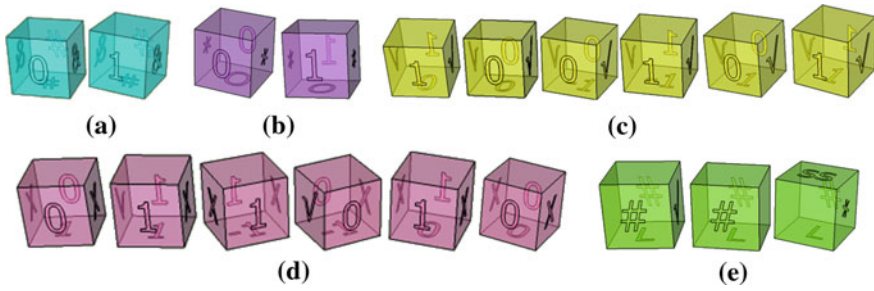
Fig. 3 **a** Input tiles. **b** Output tiles. **c** Pass tiles. **d** Check tiles. **e** Boundary tiles

subtypes of pass tiles in Fig. 3c, a $(\sigma_x = \sigma_{-x} = \sqrt{}; \sigma_y = \sigma_{-y} = \{0, 1\};$ $\sigma_z =$ null; $\sigma_{-z} = \{0, 1, -1\})$.

Check tiles: Check tiles are used to judge if the values of variable can make the clauses succeed. So check tiles were showed in Fig. 3d like this: $(\sigma_x = \times; \sigma_{-x} = \sqrt{}; \sigma_y = \sigma_{-y} = 1; \sigma_z =$ null; $\sigma_{-z} = 1)$, $(\sigma_x = \sigma_{-x} = \times; \sigma_y = \sigma_{-y} = 0; \sigma_z =$ null; $\sigma_{-z} = 1)$, $(\sigma_x = \times; \sigma_{-x} = \sqrt{}; \sigma_y = \sigma_{-y} = 0; \sigma_z =$ null; $\sigma_{-z} = -1)$, $(\sigma_x = \sigma_{-x} = \times; \sigma_y = \sigma_{-y} = 1; \sigma_z =$ null; $\sigma_{-z} = -1)$, $(\sigma_x = \sigma_{-x} = \times; \sigma_y = \sigma_{-y} = 0; \sigma_z =$ null; $\sigma_{-z} = 0)$, $(\sigma_x = \sigma_{-x} = \times; \sigma_y = \sigma_{-y} = 1; \sigma_z =$ null; $\sigma_{-z} = 0)$.

Boundary tiles: To control the growing direction of self-assembly and make the assembling smoothly, we designed the boundary tiles in Fig. 3e, just liked this: $(\sigma_y = \sigma_{-y} = \#; \sigma_{-z} = L), \quad (\sigma_x = \sqrt{}; \sigma_y = \sigma_{-y} = \#; \sigma_{-z} = L), \quad (\sigma_x = ^*; \sigma_{-y} = \#; \sigma_z = ss; \sigma_{-z} = L)$.

3 3D DNA Self-Assembly for SAT Problem

In order to make the assembling more organized, we introduced a mechanism to instruct the assembling. A strength function g: $\sum \times \sum \rightarrow R$ is considered such that mismatched surfaces have no interaction strength and matching surfaces have positive strengths. A tile may be added to an assembly if the summed strength of its interactions with neighbors exceeds a threshold called anneal temperature. In here the variable is a constant "3". Thus, only tiles can be assembled when they match up to three surfaces in a position.

In this section, we focused on the SAT problem and did some explorations on the algorithmic model. Firstly, we introduced how the 3D tiles to judge a Boolean SAT formal is correct. Secondly, the complex of this algorithm was explained.
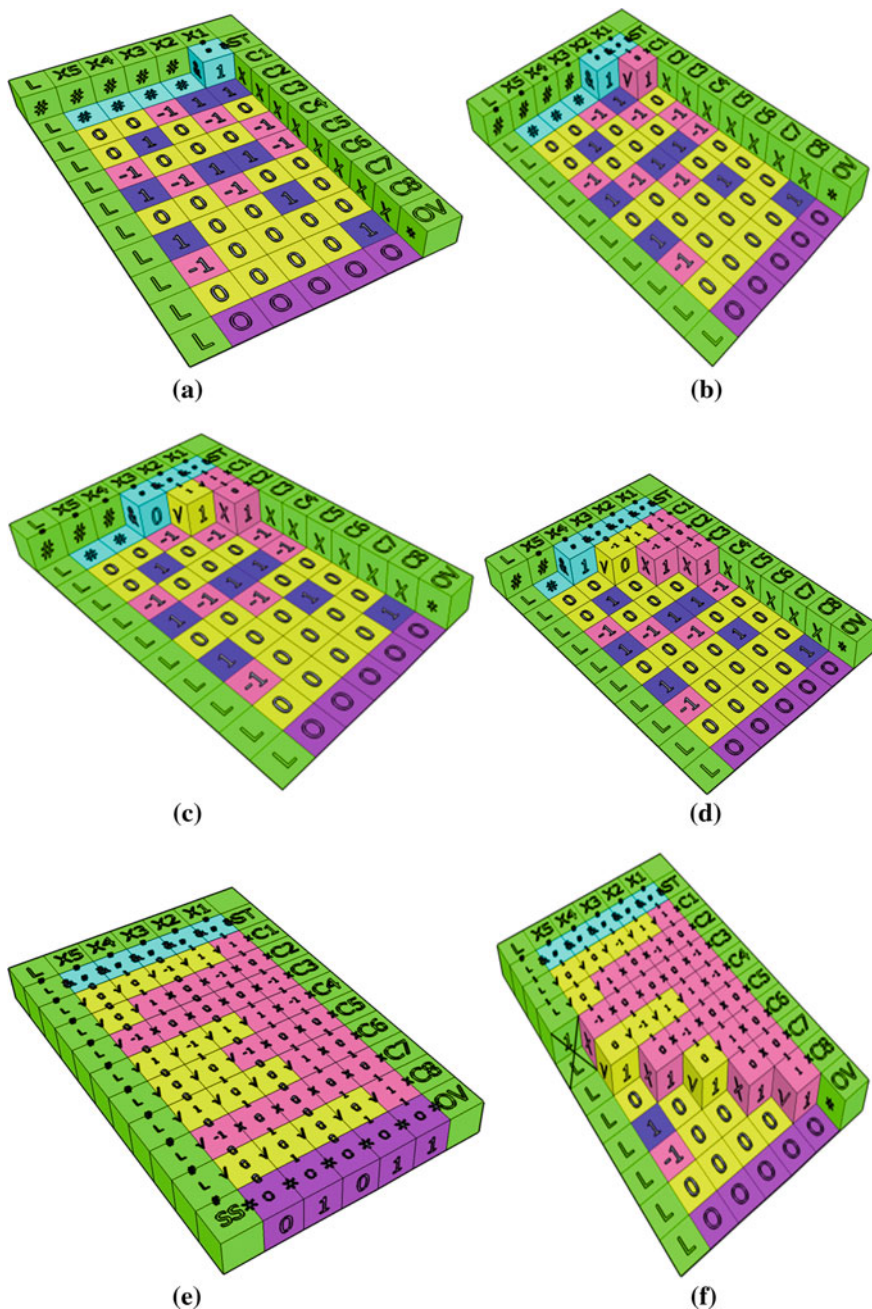
Fig. 4 (a) The first step of assembling, the first input tile was assembled. (b) The second step of assembling. (c) The third step of assembling. (d) The fourth step of assembling. (e) The end of the assembling, a successful self-assembly with proper vertex assembly was presented. (f) An unsuccessful self-assembly

3.1 The Successful Self-Assembly and Unsuccessful Self-Assembly

The growth of self-assembly follows the flow of the nondeterministic algorithm. The random set of variables (at here we assumed they were 11010) would be checked in this section and each process can be completed in a unit time during assembling. Figure 4 showed the details of each steps.

Figure 4a showed the first step of assembly, the first variable value would be assembled, which was designed as the input tile. The details were: input tile $(\sigma_x = \sigma_{-x} = \&; \sigma_y = 1; \sigma_{-y} = \#; \sigma_z = \text{null}; \sigma_{-z} = \#)$ was attached to the first "#" type from right which represented x_1, with "#", "&" and "#" matched on the deferent surface.

The second step: Input tile $(\sigma_x = \sigma_{-x} = \&; \sigma_y = 1; \sigma_{-y} = \#; \sigma_z = \text{null}; \sigma_{-z} = \#)$ and check tile $(\sigma_x = \times; \sigma_{-x} = \sqrt{}; \sigma_y = \sigma_{-y} = 1; \sigma_z = \text{null}; \sigma_{-z} = 1)$ were attached the relevant positions with the sticky ends (Fig. 4b showed the second step of assembly).

The third step: In this process there were three tiles would be assembly in a unite time: input tile $(\sigma_x = \sigma_{-x} = \&; \sigma_y = 0; \sigma_{-y} = \#; \sigma_z = \text{null}; \sigma_{-z} = \#)$, pass tile $(\sigma_x = \sigma_{-x} = \sqrt{}; \sigma_y = \sigma_{-y} = 1; \sigma_z = \text{null}; \sigma_{-z} = 1)$ and check tile $(\sigma_x = \sigma_{-x} = \times; \sigma_{-y} = \sigma_y = 1; \sigma_z = \text{null}; \sigma_{-z} = 0)$ (Fig. 4c showed the details of this step).

The fourth step: there were four tiles assembled in a unit time: input tile $(\sigma_x = \sigma_{-x} = \&; \sigma_y = 1; \sigma_{-y} = \#; \sigma_z = \text{null}; \sigma_{-z} = \#)$, pass tile $(\sigma_x = \sigma_{-x} = \sqrt{}; \sigma_y = \sigma_{-y} = 1; \sigma_z = \text{null}; \sigma_{-z} = 1)$, check tile $(\sigma_x = \times; \sigma_{-x} = \times; \sigma_{-y} = \sigma_y = 1; \sigma_z = \text{null}; \sigma_{-z} = 0)$ and $(\sigma_x = \times; \sigma_{-x} = \times; \sigma_y = \sigma_{-y} = 1; \sigma_z = \text{null}; \sigma_{-z} = 0)$ (Fig. 4d showed the details of this step).

From the above analysis, we know that the direction of self-assembly follow the oblique line. So when the last tile $(\sigma x = *; \sigma - y = \#; \sigma - z = \text{SS}; \sigma - z = \text{L})$ is attached to the surfaces labeled with "E", "*" , "#" which means the accomplishment of the entire self-assembly. To see the process of self-assembly clearly, we mark the sign of the two stuck sides on their boundary line and the details are shown in Fig. 4e. So we can conclude that "11010" can make the formula succeed.

However, not all the sets of variables can make the all clauses succeed; actually, a majority of them would bring about wrong results and the improper variable assembly can be detected at a certain step of the assembling. Figure 4f presented a situation of unsuccessful self-assembly. As we can see that at the position a big black "×" marked on the top, which meant there was no suitable tile can be assembled. So, the variable set "11111" cannot make the SAT formula succeed.

3.2 Complexity Analysis

The complexity of our 3D self-assembly model is considered in terms of computation time, computation space and the number of distinct tiles [13].

Computation Time: According to the growing trend of the assembly, the computation time can be compute as follows:

$$T = \theta(n + m + 2) = \theta(n + m) = \theta(n)$$

n stands for the number of variables; m stands for the number of clause.

Computation Space: The space taken for each assembly is the volume of the assembly, which is easy to compute as follows:

$$S = \theta((n + 1)^*(m + 2)) = \theta(n^*m) = \theta(n^2).$$

Number of Distinct Tiles: The types of distinct tiles include all the subtypes in our design. Totally, there are two subtypes of input tile, two subtypes of output tile, six subtypes of pass tile, six subtypes of check tile, and three subtypes of boundary tile. In sum, it is 9.

4 Conclusion

Benefit from the high memory density and massive parallelism, DNA computing is provided with unexampled dominance in solving NP-complete problems. Concerning the SAT problem, we designed the seed configuration through a point that a variable has three forms in a clause and used the 3D DNA self-assembly to judge which sets of random variables were suitable for the formula. Our design made an effort in two questions of creating shapes and computing functions, the computation time is $\theta(n)$ and the computation space is $\theta(n^2)$. While the 3D self-assembly model augments the computational power, it also brings tough challenges to biochemical technique. So far, we cannot valid our model experimentally, that is what we are going to work on in the future.

References

1. Wang H (1961) Proving theorems by pattern recognition I. Bell Syst Tech J 40:1–42
2. Adleman L (1994) Molecular computation of solutions to combinatorial problems. Science 266:1021–1024

 3. Winfree E, Eng T, Rozenberg G (2001) String tile models for DNA computing by self-assembly. Lect Notes Comput Sci 2054:63–88
 4. Mao C, LaBean TH, Reif JH, Seeman NC (2000) Logical computation using algorithmic self-assembly of DNA triple-crossover molecules. Nature 407:493–496
 5. Sakamoto K, Gouzu H, Komiya K (2000) Molecular computation by DNA hairpin formation. Science 288(5469):1223–1226
 6. Braich RS, Chelyapov N, Johnson CR et al (2002) Solution of a 20-variable 3-SAT problem on a DNA computer. Science 296:499–502
 7. Adleman L (2000) Towards a mathematical theory of self-assembly. Technical report, Department of Computer Science, University of Southern California
 8. Brun Y (2007) Arithmetic computation in the tile assembly model: addition and multiplication. Theoret Comput Sci 378(1):17–31
 9. Brun Y (2008) Nondeterministic polynomial time factoring in the tile assembly model. Theoret Compute Sci 395(1):3–23
10. Brun Y (2008) Solving satisfiability in the tile assembly model with a constant-size tileset. J Algorithms 63:151–166
11. Jonoska N, Karl S, Saito M (1999) Three dimensional DNA structures in computing. Bio Syst 52:143–153
12. Minqi L, Jin X et al (2010) 3D DNA self-assembly model for graph vertex coloring. J Comput Theor Nanosci 7:246–253
13. Rothemund PWK, Winfree E (2000) The program-size complexity of self-assembled squares (extended abstract). In: Proceedings of the 32nd annual ACM symposium on the theory of computing. Portland, pp 459–468

Study of Bookkeeping in the Small and Medium-Sized Enterprise

Qingping Li

Abstract With the development of small and medium enterprises, there are some new accounting services, one of which is bookkeeping. But the bookkeeping has also its disadvantages. To solve the problem of many ignored questions of bookkeeping, the method of searching for countermeasures is given in this paper. The theoretical analysis and specific example show that the method is very effective.

Keywords Bookkeeping · Ignored questions · Countermeasures

1 Introduction

Recently the small and medium enterprises have been receiving significant attention in the media and blogosphere [1]. Bookkeeping refers to a series of accounting activities in which an intermediary organization approved the establishment of bookkeeping agency to business, accepts the commission of an independent accounting subject to record, account, publish. Bookkeeping agency are placed favor and accepted by small and medium-sized enterprise, but we should pay attention to its advantages and disadvantages in the progress [2].

Based on the above background, our approach starts from a motive example, leads to the necessity of solving the problem of many ignored questions of bookkeeping. The remainder is organized as follows. Section 2 gives motive example as the basis of the study. In Sect. 3, we provide a method for searching for countermeasures. Finally, we conclude in Sect. 4.

Q. Li (✉)
Department of Economics and Management, Huaian Normal University,
Huaian 232001, China
e-mail: Lqpzsq@163.com

2 Motivation Example

This section describes the Hefei City as a motivation example [3].

By the end of 2008, 48,400 enterprises have registered in Hefei City. Small medium and small enterprises are up to 43,300, accounting for 89.55 % of the total number of enterprises. According to the survey found that these small and medium-sized enterprise lack competition in the market, financial revenues and expenditures are simple, can rarely equipped with full-time accounting personnel, while bookkeeping agency institution won some favor of the small and medium-sized enterprises with its professional and comprehensive. But in small and medium-sized enterprise agent accounting practice still have many problems.

1. Profession staff shortage, bookkeeping agency quantity and scale can not meet the needs of small and medium-sized enterprises.

According to the statistics by the end of 2008, there are 17,062 men has obtained accounting qualification in Hefei. If each enterprise is equipped with a accountant, the accounting personnel notch percent will up to 64.75 %. If each large enterprise is equipped with three accounting personnel, the extent of the gap will surprisingly up to 85.64 %. At the same time, a large quantity of outstanding accounting talents poured into highly-paid large enterprises, in this accounting talent supply and demand imbalance conditions, many small and medium-sized enterprise inevitably chose bookkeeping agency. In this paper I make a statistics of specialized agency bookkeeping institution quantity and distribution of employees in Hefei, let's see the table below. Table 1 as follows shows the bookkeeping agency institutions and professional personnel quantity distribution of Hefei.

The above data display: 1, the number of Feixi county and economic and technical development zone of bookkeeping agency are less than three, and high

Table 1 *Data sources* the provincial department of finance related statistics

County, area name	Number of bookkeeping agency company	Number of employees
Yaohai area	7	25
Luyang area	36	136
Shushan area	13	45
Baohe area	10	32
High and new technology industrial development area	0	0
Economic and technological development area	1	4
New railway comprehensive test site	7	24
Feidong County	0	0
Feixi County	2	7
Changfeng County	0	0
Total	76	273

and new technology industrial development zone, However there is no agency bookkeeping agency in Changfeng County and Feidong County still now. Relatively speaking these area of small and medium-sized enterprise are more concentrated, the local existing quantity of bookkeeping agency are seriously insufficient, can not meet the needs of small and medium-sized enterprises; 2, bookkeeping agency are on a smaller scale, equipped with less professional personnel, the ability of resisting risk is limited. 76 bookkeeping agencies, only one bookkeeping agency employs 4 people, the rest of the employees are less than 4 people. The limitations of this size makes the bookkeeping agencies which engaged in the bookkeeping agency have no obvious advantage when compared to business and social individual part-time accounting, at the same time it also cuts the ability of bookkeeping agencies to take the corresponding responsibility.

2. Agency personnel quality is not good enough. The authors made a statistic on the employees of Hefei bookkeeping agency as follows:

The datas of Figs. 1 and 2 are driven from the provincial department of finance related statistics. According to it we can see that agency lacked a group of highly educated, high title of talents, uneven quality cannot guarantee the industry business level. In addition, in the actual business operation, many employees professional attitude is terrible. They often recognize the duty as a main task, and do not place emphasis on the accounting processing; bookkeeping agency is essentially tax agent. At the same time the original vouchers audit are too formal, the release of accounting information exists data distortion.

3. The cognitive degree of bookkeeping agency business is not enough

First of all, from the degree of valuation of the government department, bookkeeping agency can't enjoy the government's preferential tax and financial aspects of policy support as other industries, and even some of the county government does not support the existence of bookkeeping agency, to certain extent corresponding development. Secondly, the small and medium-sized enterprises

Fig. 1 Employees' degree structure schematic

- Undergraduate
- Higher school
- Middle-level and junior

Fig. 2 Employees' degree structure schematic

- No title
- Senior title
- Intermediate title
- Primary title
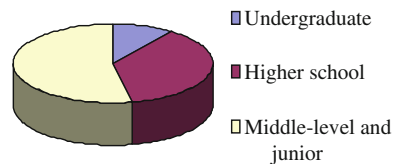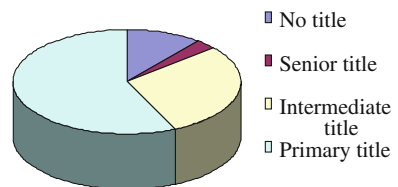
lack of awareness. Part of the small and medium-sized enterprises think book-keeping agency just help enterprise complete accounting and tax reporting, and they seldom know other services. There are some small and medium-sized enterprises, the accounting processing and tax work are often completed by part-time accounting personnel or informal agency bookkeeping personnel introduced by acquaintances. They are not willing to try new accounting services projects, which make accounting market in disordered manageable state.

4. The relevant government departments lack of supervision to bookkeeping agency

Financial department functions as bookkeeping agency for examination and approval department, every year it make routine inspection for agency book-keeping, but various daily affairs make its management often stay in bookkeeping agency license and inspection of audit. However bookkeeping agency work process, service quality, have no effective supervision. The rules and regulations, internal management, business operation and so on cannot implement necessary guidance either. The tax authorities believe that since the financial department is its examination and approval department, the business which it mainly engaged in is the financial accounting, bookkeeping agency institutions as long as the tax can. The ministry of commerce and only to the agency bookkeeping institution is legal for a single operation. This makes the bookkeeping agency institutions lack effective supervision and even supervision vacancy. The poor supervision leads to the disorder of bookkeeping agency market. Some employees have not obtained even accountant from job seniority card; some agency bookkeeping institution shall entrust the requirement of the unit for the issue of false financial information, wantonly evade taxes. Some individuals contract business personally in low price; the formal agency bookkeeping institution but accepts less business. Bookkeeping agency industry markets appeared on the phenomenon of vicious competition, appear even a lot of illegal groups without business license and permitted book-keeping agency.

5. Agent in charge to the trustee legal status has not been completely eliminated

In daily bookkeeping agency business process, bookkeeping agency company will inevitably encounter some of this "the customer", they only let bookkeeping agency institution remember tax "external account", provide documents or income material violating the fact, there is deliberately withheld income, take more expenses tax evasion behavior, while bookkeeping agency company's accounting personnel and don't know the actual situation of agency enterprises, but according to the provided enterprise the material, enter an item in an account. If leakage, deferred tax violation of tax regulation, the two sides will usually buck-passing, cause the dispute. Therefore, the trustee legal status problem has not been completely eliminated.

3 The Perfect Countermeasure of Bookkeeping Business in Small and Medium-Sized Enterprise

In this section, we will provide perfect countermeasure of bookkeeping business in small and medium-sized enterprise.

1. Support bookkeeping agency industry development

First of all publicize the positive meaning and important role of bookkeeping agency positive meaning and important role[4], change the misunderstand of bookkeeping agency in some small and medium-sized enterprise, individual and industrial and commercial door and other small economic organization for bookkeeping agency misunderstandings, realize bookkeeping agency for enterprise benefits, lead more small and medium-sized enterprises and other properties of the accounting entity and to entrust implementation bookkeeping agency ranks. In addition, strengthen government support, for those bookkeeping institutions in their initial, government management department can give appropriate tax concessions to strengthen the bookkeeping agency team to make the quantity and scale of professional agency bookkeeping institution can adapt to the needs of small and medium-sized enterprises.

2. To strengthen the supervision of the bookkeeping agency institutions

As the bookkeeping agency industry competent department, financial department should active joint national tax; land tax, industry and commerce departments to implement thoroughly "bookkeeping agency management method". Supervision department should carry out bookkeeping agency business agencies access system [5]. At the same time, to all accounts of professional personnel and institutions, establish corresponding archival establish. A regular inspection and the annual inspection system, and assess credit rating, according to inspection; for those not passing the inspection undo and withdraw their bookkeeping agency license. Resolutely ban unlicensed operation, illegal operation, accounting fraud and irregularities.

3. The small and medium-sized enterprise should choose reasonable bookkeeping agency to promote the common development of both sides

Growing small and medium-sized enterprise will be the promoting of bookkeeping agency industry development. The competent departments of the enterprise should supervise the small and medium-sized enterprise whether to establish objective complete account or not. When enterprise choose bookkeeping agency, should be from the enterprise long-term reputation and interests, and resolutely put an end to hire individuals generation accounts and built deceptive accounting behavior, but should choose a suitable quality agency bookkeeping institution and account. First of all, we should check agency business license and agent certificate, in order to ensure the legitimacy of generation accounts; secondly, should investigate agency scale and professional personnel to ensure that issued by the high

quality of accounting information. Only in this way can we promote bookkeeping agency industry and small and medium-sized enterprise's common development.

4. Strengthen the agency self-building, creating the brand service company

First of all, the existing bookkeeping agency mechanism internal control system and the specific business operation specifications should further perfect. From the enterprise getting original vouchers to agency making accounting vouchers, books and prepare financial statements, file storage, etc., are involved in many documents and data transfer, personnel responsibility determination, agent fees, etc., we should establish a standardized business processes to ensure and improve the quality of accounting information.

Second, we should improve the agency bookkeeping agency employees business ability and professional moral level. Using a variety of ways and approaches to organize the employees to study training, making them know more new business, new knowledge and related laws, regulations and systems to enhance their professional ability and professional moral level. Introduce talents, strive to attract a group of highly educated high title business personnel, to improve the agency bookkeeping agency staff team construction.

Finally, bookkeeping agency institutions should strengthen the service consciousness. Generation accounts accounting personnel in addition to accounting treatment in the main work outside still can carry out for small and medium-sized enterprises to provide the cash receipts and cash payments, invoice issue, payment collection and other services. Besides, according to the mastery of the accounting information, they should analyze the financial situation entrust unite, and timely put forward the corresponding improvement Suggestions. So as to continuously improve the bookkeeping agency work level and the quality of accounting information. In the constant development and expansion process, bookkeeping agency institutions should make great efforts to make its own brand, to expand the social influence of the accounting service industry.

5. Bookkeeping agency in the process of the legal status of the trustee shall be clearly as soon as possible

Bookkeeping agency focused on simple accounting treatment work, the entrusted party can only provide original vouchers for the clients to form rather than substantial audit, the more impossible to field to check account real work. This requires in the process of bookkeeping agency, that entrust units provide original documents timely when handing over formalities, and that both parties by contracting form clear their legal responsibility to avoid unnecessary disputes. Once financial and accounting reports and related material appear distortion, such as the enterprise tax evasion, deferred tax other ACTS violating the provisions, the two sides responsibility should be clear: 1. If the client provides false accounting information caused by the original, shall be borne by the client legal liability; 2. If it is the fault of the entrusted party caused by, bookkeeping agency institutions should take the responsibility, and compensate the principal loss of client.

4 Summaries

As long as the bookkeeping agency organization strengthen its self-building, improve internal management, perfect the bookkeeping agency procedures; at the same time small and medium-sized enterprise reasonable select bookkeeping agency, the competent departments of industry and commerce strengthen the corporation with the Tax department. The bookkeeping agency industry will become our country small and medium-sized enterprise long-term cooperation partners; both will realize win–win strategy.

References

1. Yu C (2009) Analysis the accounting agency industry development present situation and the countermeasures. Econ Res Guide 22:77–78
2. Mingqiu H (2007) Enterprise agent bookkeeping analysis. J Collect Econ 2:158
3. Changfu H, Jinyan L, Min L et al (2007) My viewpoint of the development of bookkeeping agency business. Financ Issue 7:79
4. Pengyun G (2009) At present bookkeeping agency industry existing problems and countermeasure thought in our country. J Gen Account 3:102–103
5. Jianling X (2009) Bookkeeping agency industry problem and the countermeasure analysis. J Mod Trade Ind 10:165–166

Towards the Development of a Framework of Organizational Performance Integrating Information Resource Allocation and Organizational Change in China

Fagang Hu and Yuandong Cheng

Abstract The paper focused on the relationship between information resources allocation, organizational change and organizational performance. Data were gathered from large and medium enterprises in Shan Dong, China. Method of snowball was used to issue and retrieve questionnaires. Mean score was used to present the current status of large and medium enterprises of information resources allocation, organizational change and organizational performance. Correlation analysis, regression analysis were used to identify the correlation between the three variables.

Keywords Organizational performance · Information resource allocation · Organizational change

1 Introduction

The generation of computer in the middle of the 20th century aroused a new industry and technology revolution which was called the third industrial revolution or information revolution. Now in the 21st century, computer and IT (information technology) has a further development. Information has become one of the most important production factors. Therefore, the 21st century is called the information era. IT is more and more penetrating into the world of production and life of human being. Just like steam engine, railway and electricity in the previous

F. Hu (✉)
SuZhou Unversity, Suzhou, Anhui 234000, China
e-mail: hufg@163.com

Y. Cheng
Anhui University of Science and Technology, Huainan, Anhui 232001, China
e-mail: andoncheng@foxmail.com

industrial revolution, it changed organizational form, people's life style and communication way. Many researchers have been aware that information technology was different from the previous technology. It could effectively control all the symbols used in all work. It had potential that it could affect the coordinating, production, and decision-making process inside or across the organization.

This paper mainly researched on the relationships between information resource allocation, organizational change, and organizational performance of selected large and medium enterprises in Shan Dong, China. A review of related literature is necessary for the study. Researches of information resources allocation, organizational change, and the relationship between information resources, organizational change and organizational performance will be reviewed in the following.

1.1 Information Resources Allocation

Mangarelli and Klein [1] divided four basic types of information technologies being applied in enterprises which were: automation, informatization, in-depth application and communication application. Automation was a technology of manpower replacement. It automatized information processing by data processing technology, which was once done by secretarial staff. It was similar to realizing automation of factories by machines. Informatization type was to improve the ability of information processing by strengthening the ability of data editing, analyzing and presenting. In In-depth Application type, integrated circuit chip drastically improved the functions while bringing down the price. Therefore, lots of demands of using micro processor control to replace electromechanical and mechanical control. Such application opened up a new channel of data collection, transmission and application. In communication application type, sharing ability was improved. Network technologies such as E-mail and voice mail eliminated time and distance barriers. Feng and Feng [2] do study on information resource with "Entropy".

1.2 Organizational Change

Organizational change has a famous trilogy theory which is "unfreeze-change-freeze". Unfreezing mainly referred to the change of members' thought of the organization, communicational model of the organization and so on. It was the preparatory stage of organizational change. Changing mainly referred to organizational structure change and changes in strategies. Organizational structure change was the main part of the organizational change. Freezing mainly meant institutionalization of the fruits of the change. Put them into organization rules and organizational culture. Force Field Analysis was used in this theory to explain the causes of organizational change. Kurt lewin's organizational change model laid a

foundation of organizational change research. Amburgey [3] studied on organization with Dynamics.

1.3 On Relationship of Information Resources Allocation and Organizational Change

From the practice in recent years, the introduction of the information technology will not necessarily realize "Structure Optimization" and "Process Reengineering and Optimization". In some successful cases such as Golden Gate Project of the Chinese Custom and Golden Shield Project of Chinese public security department, effective application of information technology have made the organizational structure more flat, and realized the optimization of management process. Therefore, in some other practice, the application of information technology in public departments caused very limited influence. The practice shows that, the influence of information technology to organizational operation may not be the simple linear decision relationship. The internal mechanism of the process may be very complicated, which promotes the researchers to do profound analysis to the mechanism of information technology influencing organizational operation.

In the late 1990s, some researchers [4] put forward the view point of "Mutual Structure" between information resources allocation and organization. According to this view point, information resources allocation made adjustment to the rigid structure of organization with its technical characteristic, and the organization (and its internal actors) amended and reformed the technical characteristic of information technology. Finally, they presented a kind of consistency in the process of "Mutual Structure". Wang [5] and Sun [6] did empirical study on relationship between informatization investment and enterprise performance. Li [7] studied on influence of information technology on organization.

2 Presentations, Analysis and Interpretation of Data

2.1 Current Status of the Enterprises in Shan Dong

Shan Dong is relatively well developed in China. Its GDP is the third largest. In 2011, GDP of Shan Dong reached ¥ 4,542.921 billion (Guang Dong is ¥ 526.359 billion, Jiang Su is ¥ 4,860.426). Information resources allocation is also relatively well done. In large and medium-sized enterprises, more than 95 % of them have realized the connection of intranet and internet. Information human resource is rich. Training and education in information science is also in good status. About organizational change, according to the data of 2006, more than

90 % of state-owned enterprises had already finished the restructuring. Specifically, the status of information resources allocation, organizational change, and organizational performance are presented and analyzed according to the data gathered in this study.

2.1.1 Status of Information Resources Allocation

Table 1 shows the status of information resources allocation of large and medium enterprises in Shan Dong, China.

The status of information resources allocation of large and medium enterprises in Shan Dong Province is moderate level with the mean score of 3.34.

Status of Organizational Change

Table 2 shows the current status of Organizational Change of large and medium enterprises in Shan Dong, China.

The status of organizational change of large and medium enterprises in Shan Dong province is high level with the mean score of 3.60.

Status of Organizational Performance

Table 3 shows the current status of Organizational Performance of large and medium enterprises in Shan Dong, China.

Based on the computed means of gathered data, items and variables are analyzed and interpreted in order to present the current status of organizational performance of large and medium enterprises in Shan Dong Province, China.

The results show that large and medium enterprises in Shan Dong Province have good performance in terms of the three financial aspects. This also proves that Shan Dong is a relatively well developed province in China with GDP as the third largest.

The status of organizational performance of large and medium enterprises in Shan Dong province is high level with the mean score of 3.47.

Table 1 Status of information resources allocation

Item	Mean	Std.dev.	Interpretation
Information technology	3.50	0.710	High
Information content	3.38	0.679	Moderate
Information human resource	3.14	0.814	Moderate
Information resources allocation	3.34	0.625	Moderate

Table 2 Status of organizational change

Item	Mean	Std. Dev.	Interpretation
Organizational change in management	3.55	0.888	High
Organizational change in process	3.66	0.920	High
Organizational change	3.60	0.871	High

Table 3 Status of organizational performance

Item	Mean	Std. Dev.	Interpretation
Organization financial performance	3.47	0.860	High
Organization operation performance	3.52	0.782	High
Organization human resource performance	3.42	0.781	High
Organization performance	3.47	0.681	High

2.2 Analysis of Relationships

Basing on the gathered data in this study, the relationships between information resources allocation and organizational change, information resources allocation and organizational performance, organizational change and organizational performance, information resources allocation, organizational change and organizational performance will be analyzed and presented in the following.

2.2.1 Relationship Between Information Resource Allocation and Organizational Change

In order to specifically analyze the relationship between variables of information resource allocation and organizational change, correlation analysis of variables of the two is done. The results of correlation analysis with SPSS16.0 are as following (Table 4):

The results show that in large and medium enterprises, information resources allocation can positively affect organizational change. Or it can be inferred that if large and medium enterprises want to achieve successful organizational change, effective information resources allocation is important.

Table 4 Correlations of information resources allocation and organizational change

Variables	IC	IT	IHR	OCM	OCP
IT	0.690[**]				
IHR	0.528[**]	0.548[**]			
OCM	0.436[**]	0.518[**]	0.541[**]		
OCP	0.466[**]	0.588[**]	0.555[**]	0.856[**]	
OC	0.465[**]	0.567[**]	0.566[**]	0.976[**]	0.948[**]

[**] Correlation is significant at the 0.01 level (2-tailed)

2.2.2 Relationship Between Information Resources Allocation and Organizational Performance

In order to analyze the relationship between the variables of information resource allocation and organizational performance, correlation analysis of variables of the two is done. The results of correlation analysis with SPSS16.0 are shown in Table 5.

Shan Dong is one of the Provinces with high popularizing rate of information technology. From the current status of information resources allocation and organizational performance presented in this study, and combining with the statistical results here, it can be inferred that good status of information resources allocation is an important reason why large and medium enterprises in Shan Dong get good status of organizational performance.

2.2.3 Relationship Between Organizational Change and Organizational Performance

In order to analyze the relationship between the variables of organizational change and organizational performance, correlation analysis of variables of the two is done. The results of correlation analysis with SPSS16.0 are as following (Table 6):

In large and medium enterprises, management cost takes up a large proportion in the total operating expense. Organizational change reduces management hierarchies, makes it easy to communicate up-down, integrates and optimizes business processes, makes organization more flexible. All these above reduce the operating expense, and improve organizational performance in terms of market share, sales revenue, profitability and so on. These changes also improve the satisfaction of both customers and staff. Furthermore, the staff morale is improved and less staff will leave the enterprise. Combining with the status of organizational change and organizational performance of large and medium enterprises in Shan Dong Province, effective organizational change is also an important reason why they have good status of organizational performance.

Table 5 Correlations between information resources allocation and organizational performance

Variables	IC	IT	IHR	OFP	OOP	OHRP
IT	0.690^{**}					
IHR	0.528^{**}	0.548^{**}				
OFP	0.319^{**}	0.366^{**}	0.238^{*}			
OOP	0.300^{**}	0.280^{**}	0.361^{**}	0.631^{**}		
OHRP	0.344^{**}	0.320^{**}	0.386^{**}	0.484^{**}	0.585^{**}	
OP	0.380^{**}	0.383^{**}	0.386^{**}	0.847^{**}	0.872^{**}	0.810^{**}

** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

Table 6 Correlations of organizational change and organizational performance

Variables	OCM	OCP	OFP	OOP	OHRP
OCP	0.856^{**}				
OFP	0.255^{**}	0.212^{*}			
OOP	0.191^{*}	0.221^{*}	0.631^{**}		
OHRP	0.322^{**}	0.312^{**}	0.484^{**}	0.585^{**}	
OP	0.304^{**}	0.293^{**}	0.847^{**}	0.872^{**}	0.810^{**}

** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

2.2.4 Relationship of Information Resource Allocation, Organizational Change and Organizational Performance

In order to explore the mutual relationship among information resource allocation, organizational change and organizational performance, correlation analysis of variables of the three is done. The results of correlation analysis with SPSS16.0 are in Table 7.

Statistical results of relationship between total variance of information resource allocation, organizational change and organizational performance are as following:

Seen from Table 8, there are extremely significant positive correlation between information resource allocation, organizational change and organizational performance.

From the statistical results here, combining with the results and presentation above, in can be inferred that both information resources allocation and organizational change have positive correlation with organizational performance.

It can also be found out that there are significant relationship between information resource allocation, organizational change and organizational performance. And through stepwise linear regression analysis, there are both variables of information resource allocation and organizational change come into the regression equation of organizational performance. There is existence of linear.

Table 7 Correlations of information resource allocation, organizational change and organizational performance

Variable	IC	IT	IHR	OCM	OCP	OFP	OOP
IT	0.690^{**}						
IHR	0.528^{**}	0.548^{**}					
OCM	0.436^{**}	0.518^{**}	0.541^{**}				
OCP	0.466^{**}	0.588^{**}	0.555^{**}	0.856^{**}			
OFP	0.319^{**}	0.366^{**}	0.238^{*}	0.255^{**}	0.212^{*}		
OOP	0.300^{**}	0.280^{**}	0.361^{**}	0.191^{*}	0.221^{*}	0.631^{**}	
OHRP	0.344^{**}	0.320^{**}	0.386^{**}	0.322^{**}	0.312^{**}	0.484^{**}	0.585^{**}

** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

Table 8 Correlations of information resource allocation, organizational change and organizational performance (overall variables)

Variables	Information resources allocation	Organizational change
Organizational change	0.635[**]	
Organizational performance	0.451[**]	0.310[**]

[**] Correlation is significant at the 0.01 level (2-tailed)

3 Summary

The paper attempted to explore a framework that will show the interrelationship of information resources allocation, organizational change and organizational performance for large and medium enterprises in Shan Dong China. Especially, the study intended to answer the following objectives: (1) presents what is the current status of the enterprises in terms of Information Resources Allocation, organizational change, organizational performance. (2) What relationship can be established between: informational resource allocation and organizational change; information resource allocation and organizational performance; organizational change and organizational performance; information resource allocation, organizational change and organizational performance. From the analysis above, it was found that there were significant relationship between information resource allocation, organizational change and organizational performance.

References

1. Mangarelli RL, Klein MM (1994) The reengineering handbook: a step by step guide to business transformation. Am Manage Assoc 6(2):361–368
2. Feng R, Feng B (1992) Entropy. Science press, Beijing
3. Amburgey TL (1993) Resetting the clock: the dynamics of organizational change and failure. Adm Sci Q 38(1):51–73
4. Qui Z (2005) Mutual structure of technology and organization: taking application of information technology in manufacturing enterprise for example. Sociol Stud 18(2):78–80
5. Wang M (2007) Enterprise's informatization investment performance and its influencing factors: based on empirical evidence of enterprises in Zhe Jiang Province in China. Soc Sci China 7(6):81–93
6. Sun X (2010) Empirical study on relationship between Information technology investment and organizational performance. Stud Sci Sci 18(3):397–404
7. Li D (2010) Influence of information technology on organization. Mark Wkly 36(4):111–113

New Ideas for FN/RFN Queries Based Nearest Voronoi Diagram

Wencai Liu and Yuan Yuan

Abstract While most research focuses on nearest neighbor (NN) and reverse nearest neighbor (RNN) queries, the furthest neighbor (FN) and reverse furthest neighbor (RFN) query are attracting more and more attention. In this paper, we give new ideas on the FN/RFN query issue. The nearest Voronoi diagram-based (VD-based) algorithms to process FN and RFN queries on 2-dimensional location data on the fly is proposed. These algorithms are especially useful for applications which have VD as their fundamental structure. Instead of constructing a new furthest neighbor diagram (FD), we aim to fully utilize the properties of VD. Besides, we implemented our new methods as well as the most promising competing methods which have been previously proposed. The results consistently indicate that our new algorithms outperform competitors.

Keywords Nearest Voronoi diagram · Furthest Voronoi diagram · Furthest neighbor query · Reverse furthest neighbor query · Convex hull

1 Introduction

NN and RNN queries has also been well studied. Driven by the increasing number of applications, the design of algorithms to answer these queries has always been a focal point. In contrast to NN/RNN query, an FN query is to find points that have the longest distance from the given point and an RFN query is to find points that

W. Liu (✉) · Y. Yuan
College of Information Science and Technology, Jiujiang University, Jiujiang 332005 Jiangxi, China
e-mail: wensliu@126.com

Y. Yuan
e-mail: yuanyuan041031@126.com

are less influenced by a given point. Like RNN queries, RFN queries come in two flavors: monochromatic and bichromatic. The Monochromatic RFN query involves one type of data objects, while bichromatic RFN query involves datasets of two types. An RFN query issued by a restaurant on a restaurant dataset (to find the restaurants that are less influenced by the query restaurant) is a monochromatic query because both the query data and the queried data are of the same type. We can get the definitions of the furthest neighbor, MRFN query and BRFN query from [1]. In this work, we focus on monochromatic queries.

Many applications supporting NN/RNN queries naturally have the corresponding furthest versions. Therefore furthest neighbor queries and reverse furthest neighbor queries are also very important and practical. Assume there is a fire house and some chemical plants are located around it. The firemen need to know the furthest chemical plant as well as the distance from their fire house, because they must prepare enough fuel for vehicle to get to any plant as soon as possible. They never want to waste time to fill their vehicles before they start or during the driving. This is an example of FN query. Here, the group of chemical plants is data set P, the fire house is the query point p. In commercial competitions, sometimes two companies hope to merge to create a bigger firm to survive in commercial competitions. Assume there are several supermarkets in a region and we regard them as data set P. In order to improve commercial competitiveness, one of them (as the query point q) wants to find another supermarket as partner to form a combined entity. Then q may wants to search the supermarket on which q has less influence compared to other supermarkets (means the supermarket we are looking for takes q as its furthest neighbor),since for those supermarkets on which q has high influence, gobbling up is more wiser than merging. This is another example of MRFN query.

The motivation to answer FN/RFN queries using VD is largely inspired by the properties of VD and the fact that the demand for NN/RNN queries is much greater than demand for FN/RFN in an application during a period of time. In an application which is responsible for answering both of the two kinds of queries, FN/ RFN query is supporting role. With full-scale consideration, we hope tohandle both kinds of queries with the same VD structure without constructing FD.

In this paper, we provide a new train of thought to process FN/RFN query. Not that although in this paper our methods are based on pure VD structure, they can easily be adapted for most Voronoi-based index structures (VoR-Tree, VN-tree, grid-partition index and so on).

We make the following contributions in this paper.

1. We propose a new property for nearest Voronoi diagram, which inspired us to design our algorithms.
2. We propose algorithms to answer FN and RFN queries based nearest Voronoidiagram.

Table 1 provides a quick reference to the main notations.

The rest of the paper is organized as follows. Section 2 describes background and related work. Section 3 presents a property which plays an important role in

Table 1 Notions

Notation	Definition
p	The query point
q	The data set of points of interest
VD or Voronoi diagram	Nearest Voronoi diagram
dist(p, q)	The Euclidean distance between point p and point q
FN/RFN	Furthest neighbor/reverse furthest neighbor
NN/RNN	Nearest neighbor/reverse nearest neighbor
Vp	Voronoi polygon generated by point p

our next algorithms. Section 4 gives our furthest and reverse furthest neighbor query processing method. Section 5 introduces the experimental study. Finally,we conclude the paper with a brief discussion on the future work in Sect. 6.

2 Background and Related Work

In this section, we introduce definitions and notations, and survey the algorithms that are closely related to our work.

When mention FN query, we have to mention the FD, just like when mention NN query, we have to mention the VD. The definition of the FD can be found in [2]. Its almost the same as that of the VD except that closest is replaced with furthest. Given a set of distinct points in the Euclidean plane, we assign all locations in that plane to the furthest member(s) of the point set. The result is a tessellation of the plane into a set of regions associated the members in the point set. We call this tessellation the FD [2] generated by the point set. Fig. 1 shows an example of FD. FD naturally has the ability to answer FN/RFN query. To our best knowledge, there are no surveys specially conducted on FN search. In the following, we give a brief overview of the different approaches to address RFN query problem. For RFN queries, the work most closely related to this paper is the recent work [1, 3]. Two RFN query types (MRFN and BRFN) are originally defined in [1]. The authors of the paper proposed the progressive furthest cell (PFC) algorithm and the convex hull furthest cell (CHFC) algorithm to handle RFN query on the fly, both accepted R-tree index as their inputs and used the furthest Voronoi cell(Voronoi ploygon) to determine the RFN query results. CHFC algorithm has made a great improvement against PFC, it is in general an efficient and new method for solving RFN query problem while still a little expensive mainly due to the construction of convex hull and range query on R-tree. R-tree index rely on search algorithms cannot avoid backtracking. And this backtracking is usually increased by an overlapping partitioning schema. A complete algorithm shortly named PIV to handle RFN query was proposed in [3]. PIV adopted matrix index as data structure (Fig. 1) .

Fig. 1 FD
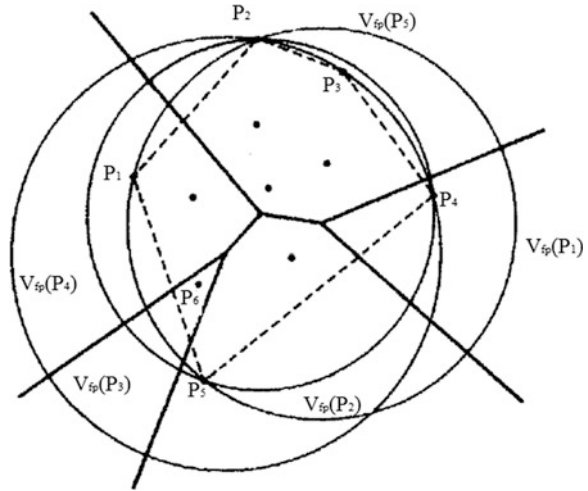
3 Preliminaries

In this section, we mainly present some important properties for VD, based on which we will propose algorithms for answering FN/RFN query later. Any Voronoi diagram has the property that related to the convex hull and can be described as follow in next property.

Property 1: For a Voronoi diagram generated by a set of distinct points P, $P = p1,\ldots, pn \in R2$ ($2 \le n < \infty$), a voronoi polygon V (Pi) is unbounded if and only if Pi is the boundary of the convex hull of P [2].

A convex hull property of RFN query is proposed in [1], which can be concluded as Property 2 and we also call it Convex-Hull Property in this paper. Property 2 (Convex-Hull Property): Only points belong to convex hull have chances to become someones furthest neighbors i.e. Only points belong to convex hull have reverse furthest neighbors.

We call it the FN property of Voronoi diagram (nearest VD). This is a new property for VD. This property is the foundation of our algorithms, which in-spires us to design the following algorithms based VD.

4 Our Solutions

4.1 FN Query Processing

So far we answer FN queries by constructing FD, just like we use nearest use Voronoi diagram to deal with NN queries. An FN query is so similar to a NN query that few even bother to ask if there are some other ways to answer FN queries except constructing FD. Its a simple approach but also time-consuming

with the space data increases. Generally speaking, the demand for NN queries is much greater than that of FN queries in an NN/RFN-FN/RFN-query-mixed application during a period of time. For this kind of application, it is in general not advisable or even not necessary to waste too much resources (such as space and time) to process FN/RFN query. Its clear that we will obtain much better performance if we can handle FN queries directly with the VD instead of constructing an additional FD.

In contrast to the existing approach where we need to construct VD, we give new ideas. We proposed an algorithm called VDFN shown in Algorithm 1 to challenge the FN queries, based on the properties mentioned before. The algorithm takes as inputs the query location point q and VD structure and returns the furthest neighbor point of q. It is important to note that the algorithm is based on VD structure.

Input: Query point q,VD vd
Output: FN query results

  Initialize candidate set $Cand \leftarrow \emptyset$, local variable $Tmp \leftarrow 0$, and result set Result$\leftarrow \emptyset$;
  Compute C_{vd} (the convex hull) with vd using VDCH algorithm, and insert C_{vd} points to Cand;
  foreach pi $\in$ Cand do
  Compute $Dis = \mathrm{dist}(q,p_i)$;
  if $Tmp < Dis$ then
    $Tmp = Dis$;

    Result$\leftarrow \emptyset$;{Remove fake points.}
    Insert p_i to Result;

  end
  if $Tmp = Dis$ then

    Insert p_i to Result;

  end
  end
  return Result;
Algorithm 1: VDFN

4.2 RFN Query Processing

The MRFN of q is a set of points from P that take q as their furthest neighbors comparing to all points in P, that is to say, MRFN(q,P) = p|p $\in$ P, fn(p,P q) = q [1].

Our algorithm, called VDRFN, to compute the result for a RFN query is shown as Algorithm 2. The algorithm takes q as inputs the query location point q and VD structure and returns the points taking q as their furthest neighbors among all points in vd.

4.3 Compute Convex Hull

The computation of convex hull is an important part of both two algorithms above. Here we present a simple but efficient way to compute the convex hull based VD structure. Our approach is called VDCH shown in Algorithm 3 to get the convex hull points based VD, which is inspired by Property 1. According to Property 1, we know that unbounded Voronoi polygons and vertices of the convex hull are in one-to-one correspondence. Since we have already had a VD structure object vd, the key is to find all unbounded Voronoi polygons from vd.

In order to find the unbounded ones, we could begin by making each polygon as being bounded or unbounded. This can be done by scanning each polygon once, in $O(n)$ (n is the total number of points in vd) time. Begin with any unbounded polygon V_{pi} and examine its edges, looking for an edge that is shared with another unbounded polygon V_{pj}. Points i and j are consecutive on the convex hull and we continue by searching V_{pj}, proceeding point by point around the hull. No Voronoi edge is examined more than a constant number of times [4].

> Input: Query point q, VD vd
> Output: RFN query results
>
> Initialize candidate set *Cand* $\leftarrow \emptyset$, local variable *Tmp* $\leftarrow 0$, and result set Result$\leftarrow \emptyset$;
> Get point set P from vd and insert them to Cand;
> Compute C_{vd} (the convex hull) with vd using VDCH algorithm;
> if q $\notin C_{vd}$ then
> Return Result;
> end
>
> foreach $o_i \in$ *Cand* do
>
> if $\exists dist_{pi} \in Cand$ (p_i, o_i) > $dist$(q, o_i) then
>
>   Discard o_i;
> end
> else
>   Insert o_i into Result;
> end
> end
> return Result;
> Algorithm 2: VDRFN
> Algorithm 3 is the method we actually take to compute convex hull in our experiment.
> Input: VD vd
> Output: Convex hull points
> Initialize result set $C_{vd} \leftarrow \emptyset$;
>
> foreach $V_{pi} \in$ Voronoipolygonsof vd do

```
    if V_pi is unbounded then
        Insert p_i to C_vd;
    end
    end
    return C_vd;
  Algorithm 3: VDCH
```

5 Experimental Evaluation

To confirm the correctness and efficiency of the proposed algorithms, we perform extensive experiments for evaluation. All the experiments are executed on an Intel-based computer and Linux OS. The CPU is Pentium(R) Dual-Core 3.06 GHz and the amount of main memory is 1 GB. The programs are implemented in C++ language, using the open-source libraries: Spatial Index Library and CGAL Library. We use voronoi-diagram-2 class from CGAL Library as VD structure.

Both synthetic and real data sets are used for experiments. The synthetic datasets include uncorrelated uniformly generated (UN) points, random-cluster distribution (RC) points and Correlated Bivariate (CB) points. The real data sets were obtained from the digital chart of the world server where points define the road networks from California, San Francisco and USA. The data sets are available online [5]. An example of the synthetic datasets is illustrated in Figs. 2, 3 and 4.

5.1 Performance of FN Searching Algorithms

Our proposed VDFN algorithm outperforms the art-of-state algorithms. The results are confirmed by both real data set and synthetic data set. The corresponding experiment is shown in Figs. 5 and 6. Figure 5 depicts the results of the experiments for algorithm GFN on both synthetic and real data sets. GFN is the kind of
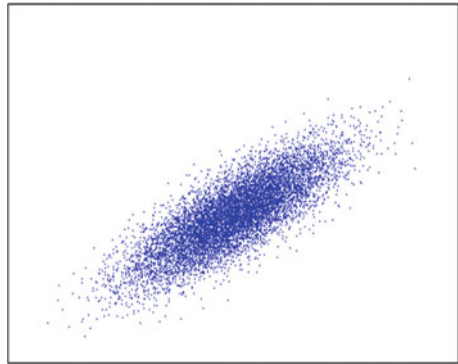
Fig. 2 Uniform distribution (UN)
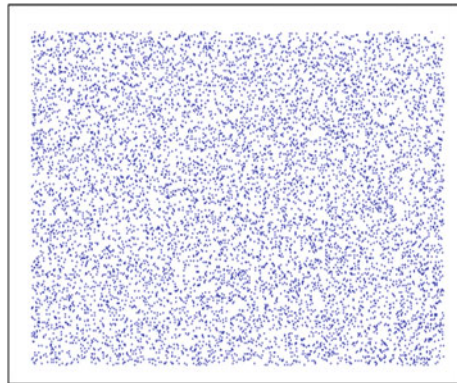
Fig. 3 Random clusters (RC)
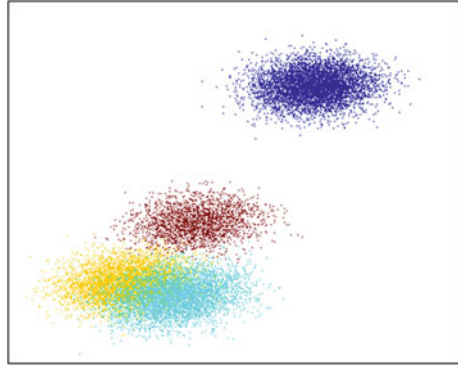

Fig. 4 Correlated bivariate (CB)

FN searching algorithm which looks for the furthest neighbor by constructing furthest Voronoi diagram. Like to find nearest neighbors, there are mainly two steps in GFN: firstly construct the furthest Voronoi diagram, secondly, find the furthest neighbor(s) based on the location of the query point.

We use the same data sets in Figs. 5 and 6. From the two figures we note that the results are not of same order of magnitude. VDFN far outperform GFN. That is because algorithm GFN spends lots of time to construct the furthest Voronoi diagram while algorithm VDFN uses the nearest Voronoi diagram directly. In addition, the VDFN algorithm performs much better than GFN by about three magnitude.

5.2 Performance of RFN Searching Algorithms

In this section, we examine the performance of VDRFN algorithm, comparing with PIV and CHFC. As illustrated in Figs. 7 and 8, the experiment results are reported for the comparison on real data and synthetic data, and the VDRFN algorithm outperforms the other two algorithms in terms of CPU cost.

**Fig. 5** Performance of GFN on different data sets
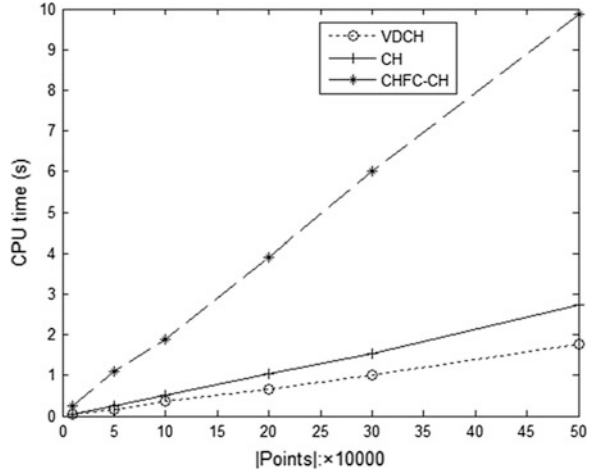
**Fig. 6** Performance of VDFN on different data sets
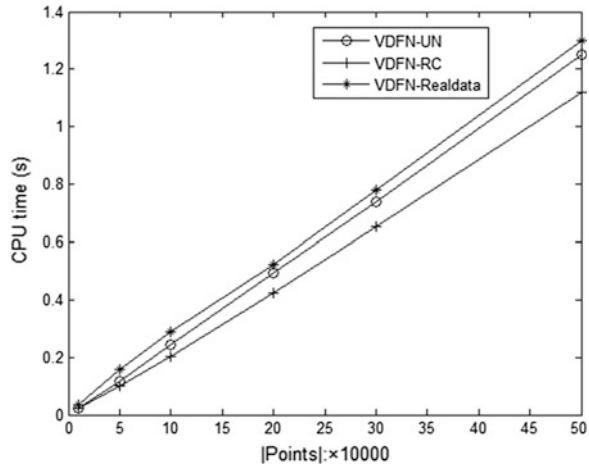
6 Conclusion and Future Work

In this paper, we proposed a new property for VD, and we call it FN Property. The property is very useful when we deal with FN/RFN queries with VD. We also proposed two algorithms called VDFN and VDRFN to answer FN and RFN queries respectively. In these algorithms, we take full advantages of the VD due to the FN Property. Without constructing a new furthest Voronoi diagram, these algorithms are efficient and save much space. We finally showed in an experimental evaluation that our approach is efficient for various kinds of data and clearly outperforms the state of the art algorithms.

Fig. 7 RFN comparison on
real data set sets
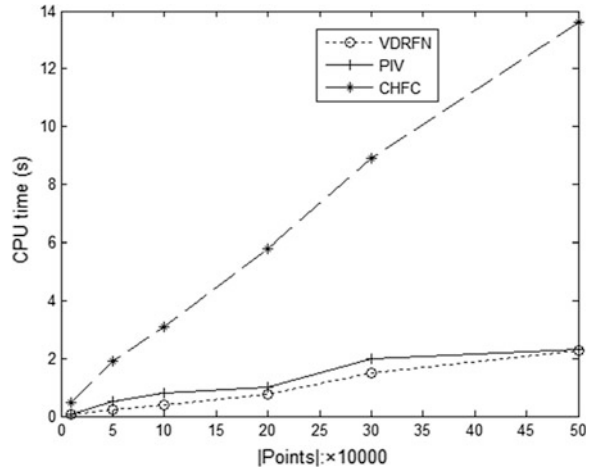
Fig. 8 RFN comparison on
synthetic data set (RC)
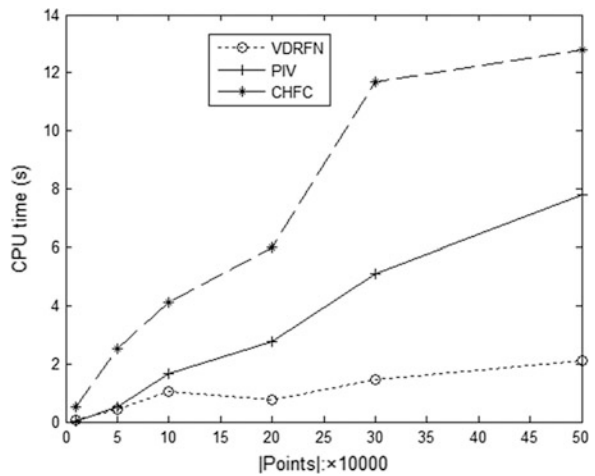
Our future research interests are focused on handling kFN/kRFN queries and
the issue of FN/RFN queries in road network. Also we are planning to put efforts to
apply this work to the real applications.

References

1. Yao B, Li F, Kumar P (2009) Reverse furthest neighbors in spatial databases. In: IEEE
   international conference on data engineering, pp 664–675
2. Okabe A, Boots B, Sugihara K (1992) Spatial tesselations concepts and applications of
   Voronoi diagrams. Wiley, Chichester

3. Liu J, Chen H, Furuse K, Kitagawa H (2010) An efficient algorithm for reverse furthest neighbors query with metric index. In: Database and expert systems applications. Springer, Berlin
4. Shamos MI, Hoey D (1975) Closest point problems. In: Proceedings of the sixteenth annual institute of electrical and electronic engineers symposium on the foundations of computer science, IEEE computer society, pp 151–162
5. www.cs.fsu.edu/ lifeifei/SpatialDataset.htm

Video Summarization by Robust Low-Rank Subspace Segmentation

Zhengzheng Tu, Dengdi Sun and Bin Luo

Abstract Video summarization provides condensed and succinct representations of the content of a video stream. A static storyboard summarization approach based on robust low-rank subspace segmentation is proposed in this paper. Firstly, video frames are represented as multi-dimensional vectors, and then embedded into a group of affine subspaces using low-rank representation according to the content similarity of the frames in the same subspace. Secondly, a series of subspaces are segmented based on the Normalized Cuts algorithm. The video summary is finally generated by choosing key frames from the significant subspaces and ranking these key frames in temporal order. The experimental results demonstrate that the proposed summarization algorithm can produce crucial key frames and effectively reduce the visual content redundancy in summary comparing with the conventional approaches.

Keywords Video summarization · Low-rank representation · Subspace segmentation · Normalized cuts

Z. Tu · B. Luo (✉)
Key Lab of Intelligent Computing and Signal Processing of Ministry of Education,
Anhui University, Hefei 230039, China
e-mail: luobinahu@yahoo.com.cn

Z. Tu
e-mail: tuzhengzheng@gmail.com

Z. Tu · D. Sun · B. Luo
School of Computer Science and Technology, Anhui University, Hefei 230601, China

Z. Tu · B. Luo
Key Lab of Industrial Image Processing and Analysis of Anhui Province,
Hefei 230039, China

1 Introduction

Rapid growth of the amount of digital video has made automatic video summarization a necessity for managing and browsing. Video summarization provides a condensed version of a full length video through searching the most important content within the stream. Video summarizations have been applied in several broad areas, such as video retrieval and interactive browsing systems [1].

Conventionally, automatic video summarization can be categorized into two major types: static storyboard [2, 3] and dynamic video skimming [4].In static storyboard, key frames are extracted and ranged in temporal order. Dynamic video skimming provides consumer a vivid visual impression by connecting the critical video section in a given time length. The key frame is a vital factor in a video summarization as a straightforward expression of video content. Most of existing key frame extracting approaches [5] are based on shot-level. The shots of the original video are first detected, and then one or more key frames are extracted for each shot. But several continuous shots usually represent the same content,thus this kind of approaches hardly summarize the content of the video without redundancy. So the prime challenge for story-board summary is to produce the key frames that represent the essence of the video content while removing redundancy among key frames as much as possible. This is the main issue we will tackle in this paper.

Video data are normally ranked with the development sequence of events or in line with the need of content representation. But we take another angle of view in this paper. We consider that video frames belong to a number of subspaces according to the content similarity of the frames in the same subspace. Based on the view, a video summarization technique by low-rank representation (LRR) and robust subspace segmentation is presented. As a compressed sensing technique, LRR represents a data vector as a linear combination of the other vectors in a subspace. Given a set of vectors generated from video frames, LRR seeks the lowest-rank representation of all data jointly, and these video data are embedded into a union of multiple linear (or affine) subspaces. Then, a series of subspaces are segmented based on Normalized Cuts algorithm. Key frames in the significant subspaces are chosen to represent the content of corresponding video segments, and video summary is generated by arranging these key frames in temporal order.

The rest of the paper is organized as follows. The low-rank subspace representation technique is presented in Sect. 2. Our proposed video summarization approach is presented in Sect. 3 followed by the comparison experiments with Open Video storyboard, MKFE, and K-means clustering in Sect. 4. Finally, the conclusion is provided in Sect. 5.

2 Low-Rank Subspace Representation

The parametric model of a given data set can characterize these data. For this purpose, linear models such as the linear subspaces are possibly the most common choice, mainly because they are easy to compute and are also often effective in real applications. The researches and approaches about the subspaces have been gaining many achievements in recent years, for example, matrix competition [6, 7] problem.

Low-rank representation is described as follows [8]:

Consider a set of data vectors $X = [x_1, x_2, \ldots, x_n]$ in R^D, each of which can be represented by the linear combination of the basis in a "dictionary" $A = [a_1, a_2, \ldots, a_m]$:

$$X = AZ \tag{1}$$

where $Z = [z_1, z_2, \ldots, z_n]$ is the coefficient matrix with each z_i being the representation of x_i. The dictionary is often over complete and hence there can be multiple feasible solutions to problem (1). It has been observed [9] that sparse representations using an appropriate dictionary A may reveal the clustering of the points x_i. However, as mentioned above, sparse representation may not capture the global structures of the data X. However, low rankness may be a more appropriate criterion by solving the following problem:

$$\min_Z rank(Z) \quad s.t., X = AZ \tag{2}$$

The optimal solution Z^* of the above problem is called the "lowest-rank representations" of data X with respect to a dictionary A. The above optimization problem is difficult to solve, so the following convex optimization provides a good surrogate for problem (2):

$$\min_Z ||Z||_* \quad s.t., X = AZ \tag{3}$$

Here, $|| \cdot ||_*$ denotes the nuclear norm of a matrix, i.e., the sum of the singular values of the matrix.

In order to segment the data vectors into their respective subspaces, an affinity matrix that encodes the pairwise affinities between data vectors is computed. We use the data set X itself as the dictionary, that is, $A = X = [x_1, x_2, \ldots, x_n]$, therefore problem (3) becomes:

$$\min_Z ||Z||_* \quad s.t., X = XZ \tag{4}$$

There always exist feasible solutions because a data vector can be used to represent itself in LRR.

Theorem 1: *Assume that the data sampling is sufficient such that $d_i=$ rank $(X_i) < n_i$. If the subspaces are independent then there exists an optimal solution Z^* to problem* (4) *that is block-diagonal:*

$$Z^* = \begin{bmatrix} Z_1^* & 0 & 0 & 0 \\ 0 & Z_2^* & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & Z_k^* \end{bmatrix}$$

with Z_i^* being an $n_i \times n_i$ matrix with $rank(Z_i^*) = d_i, \forall i$. The proof is provided in [8].

In next section,we will propose a new video summarization approach which is based on the above theory.

3 Proposed Approach for Video Summarization

First, the video frames are pre-sampled to reduce the number of frames that will be used in the next algorithm. The rationale for pre-sampling is that any video with 20–30 fps will generate too many frames with redundancies. It is possible to capture the video content with pre-sampled video frames using a judicious sampling rate. The sampling rate we use in our experiment is one out of every 5 frames. The experimental results show that the quality of the summarization is not affected by pre-sampling.

Second, each frame of the pre-sampled video is converted into luminance frame by taking a linear combination of the red, green, and blue color channels of the corresponding frame [10]. Let $y_i \in R^n, 1 \le i \le N$, be the vector representation of the ith luminance frame with pixels arranged in lexicographic order and N be the total number of frames of the input video after pre-sampling. Because the dimension of y_i is too high, we use Principal Component Analysis (PCA) to reduce dimension of the vector. The vector from dimension reduction of y_i, denoted as $x_i \in R^m, (m < n, m < N)$ is acted as a feature of the ith frame. The vector can represent the content of the ith frame much more efficiently than other traditional features such as SIFT, color histogram, etc. So the set of video data is $X = [x_1, x_2, \ldots x_N], x_i \in R^m$.

In the proposed approach, dictionary A is constructed by arranging (in temporal order) all the x_is and the ith one as the column of A, i.e. $A = X = [x_1, x_2, \ldots x_N]$. Furthermore, each column of the dictionary is normalized to one, i.e., $x_i^T x_i = 1$, where "T" denotes the matrix transpose operation. We choose to solve an optimization problem (5) [11] to estimate Z of Eq. (4)

$$\min_Z \frac{1}{2} ||XZ - X||_F^2 + \mu ||Z||_* \qquad (5)$$

where $\mu > 0$ and $|| \cdot ||_F$ denotes the Frobenius-norm.

Algorithm: Video Summarization by Low-Rank Subspace Segmentation
Input: Pre-sampled video, number of subspaces (k)

Input: Pre-sampled video, number of subspaces (k)
Output: Video summary which includes k_0 frames
STEP 1: Convert each frame into vector and reduce its dimension;
STEP 2: Obtain the low-rank representation Z by solving problem (5);
STEP 3: Construct an undirected graph by using Z to define the affinity matrix of the graph;
STEP 4: Use NCut to segment the vertices of the graph into k subspaces;
STEP 5: For k_0 larger subspaces, select the middle frame as a key frame;
STEP 6: Arrange key frames in temporal order to generate the summary.

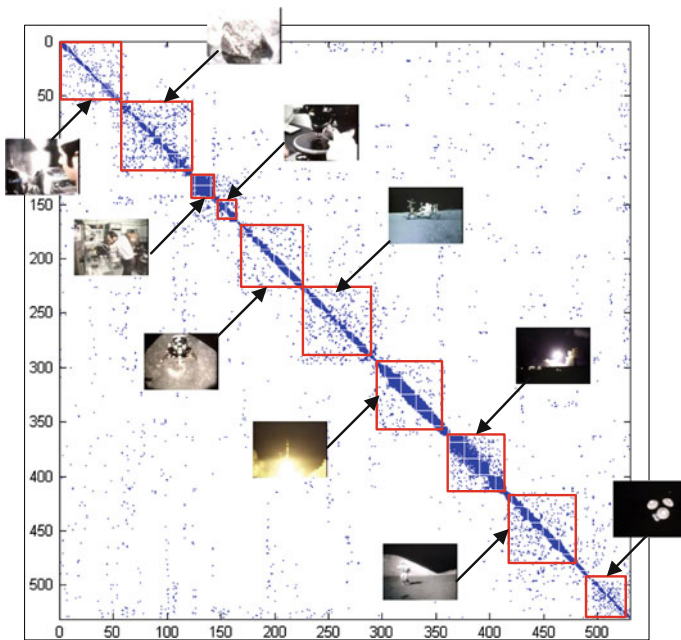**Fig. 1** Flow chart of the proposed video summarization method

Fig. 2 Coefficient matrix Z and subspaces with NCut segmentation

After estimating Z, $Z \in R^{N \times N}$, where Z is a symmetric matrix, Normalized Cuts [12] is applied on the Z to segment the subspaces. The number of clusters is set equal to the desired number of subspaces. After finding k_0 clusters which are larger, and the middle frame from each is selected as a key frame. Finally, these key frames are sorted in temporal order to produce the summary.

The main steps of the proposed summarization algorithm are as follows (Fig. 1):

Figure 2 is the coefficient matrix $Z = [z_1, z_2, \ldots, z_n]$ with z_i being the representation of x_i. As an entry of vector z_i, z_{ij} is the coefficient representing the relevance (content similarity level) of frame vector x_i and x_j. The range of z_{ij} is

from 0 to 1. From this plot, we can see the sparse representation Z is approximate to block-diagonal obviously, that is the within-cluster affinities form a connected component and the between-cluster affinities are sparse. This block structure is consistent with the content similarity of videos. In particular, non-zero coefficients with large value are aggregated in diagonal block of Z, which are emphasized using a series of red boxes. In each block, the video frame vector x_i can be represented by the linear combination of other vectors. Because of camera motion, e.g., pan, zoom, or some dynamic effects like fade-in and fade-out, there are thick diagonal lines in some blocks, such as the "Rocket Launching" segment. It indicates that these frames can be represented mostly by several adjacent frames. In fact, they will not influence the clustering result if we choose appropriate number of subspaces. Finally, we use the Normalized Cuts to segment these subspaces, and then select key frames from significant subspaces.

4 Experimental Results

A total of 50 video clips from Open Video Project [13], with resolution of 320×240 or 640×480, frame rates from 24 to 30 fps, and average duration of 50 s, are selected to validate the proposed approach. Contents of the videos vary from historical film, natural scenery, speech, to news, outdoor activities, trips, etc. Every video clip contains several shots. Our algorithm has been implemented by MATLAB R2010 on a PC with 2.4 GHz CPU and 2 GB RAM running Windows XP.

To verify the effectiveness of the proposed summarization technique, the meaningful comparison experiments with Open Video storyboard [13], MKFE [14] (camera motion-based key frame extraction) and K-means clustering were conducted. The Open Video storyboard, as a third party algorithm, is generated using the curve simplification algorithm and some manual intervention [15]. The MKFE algorithm, usually regarded as a benchmark algorithm, first segments an input video into homogeneous parts based on major types of camera motion, e.g., pan, zoom, pause, and steady, and then dedicated rules to extract key frames from each section. K-means clustering algorithm, with low computation load and reasonably good performance, also conducts the luminance frames and extracts a key frame from each clustering to represent the video content.

Due to page constraints, only two sets of results are presented in this paper. The videos shown in Figs. 3 and 4 are with duration of 53 and 56 s respectively, and both captured at 29.97 fps. From the results, we can see that, Open Video storyboards of both videos have the biggest number of key frames with 15 and 20 respectively, which in turn means it can produce more details. However, the redundancy emerges consequentially, e.g. the 9th to 12th key frame in Fig. 3a actually represent the same content, the same as the last three key frames in Fig. 3a and the first two key frames in Fig. 4a.

(a)

(b)

(c)

(d)
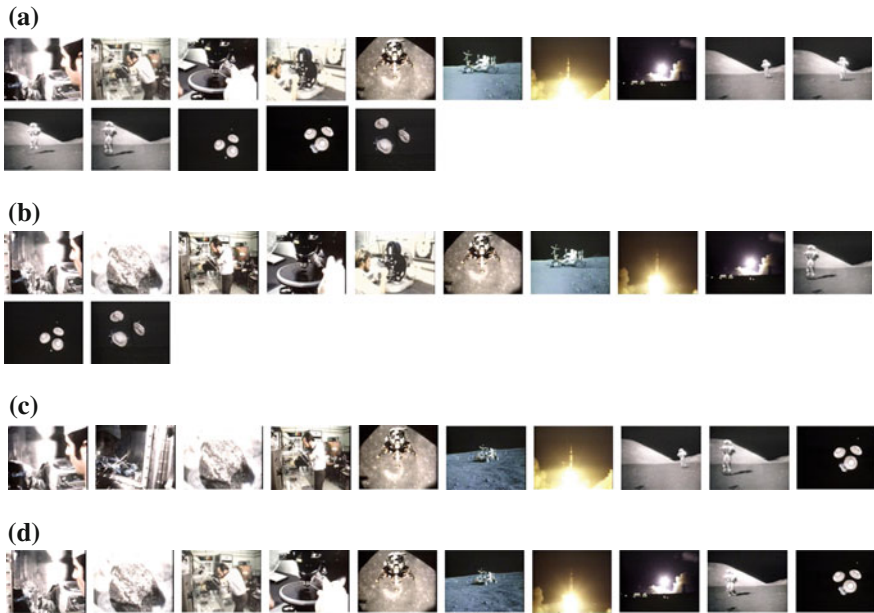
Fig. 3 Summary by (**a**) open video storyboard, (**b**) MKFE, (**c**) K-Means, and (**d**) low-rank subspace segmentation for "NASA 25th Anniversary Show, Segment 07"

The MKFE-based summarization algorithm abstracts the summary depending on major types of camera motion. But one single key frame from each shot can not reflect the essential change of visual content correctly because a long shot with pan or zoom often arises some different scenes or objects, e.g. the first shot of "New Indians, Segment 01" is a long shot, but MKFE produces only one frame, that is the first frame in Fig. 4b. Meanwhile, several continuous shots maybe describe the same content, e.g. the last two key frames in Fig. 3b. Therefore, summarization based on MKFE brings redundancy or absence of significant frame inevitably. As we can see in Fig. 3d and Fig. 4d, our approach can effectively reduce redundancy since it makes use of content similarity of frames to cluster. It is insensitive to the camera motion and shots types.

We conduct experiments based on K-means algorithm to make some comparisons. In the listed experiments, the numbers of clusters are set as 10 and 16 respectively, the same as the numbers of key frames in the proposed approach. Obviously, the results in Figs. 3c and 4c contain some redundancy, meanwhile miss lots of prime information. This is due to the instability of K-means algorithm. For example, the 9th key frame in Fig. 3c shows the same content with the 8th, and necessary content such as "microscope" has not been extracted.

In our approach, each video is grouped into a series of low-rank subspaces, and then, k_0 subspaces with more than 20 dimensions are chosen to represent this video content. A set of key frames is obtained from these subspaces. In fact, the

(a)

(b)

(c)

(d)

Fig. 4 Summary by (**a**) open video storyboard, (**b**) MKFE, (**c**) K-Means, and (**d**) low-rank subspace segmentation for "New Indians, Segment 01"

dimension of the subspace corresponds to the size of the cluster, it also reflects the length of the video segment simultaneously. We assume unimportant content usually holds shorter time duration, such as the 4th frame in Fig. 3a belongs to a segment lasting 1 s, the 5th key frame and the 7th key frame in Fig. 4a correspond to the segments lasting 2 s and 1.5 s respectively. As shown in Figs. 3d and 4d, our approach removes most visual content redundancy among key frames of summary by extracting key frames from the significant clusters.

Based on all experimental results, we conclude that 31 of the 50 videos have fewer key frames than Open Video storyboard; 9 videos have identical summaries; 10 videos show more number of key frames than the Open Video storyboard; 3 of the 50 videos show mismatched key frames with not much overlap while 2 of those 3 videos were due to the decoding problems and the last one video show very low variance after PCA, which can be easily fixed with a better implementation of PCA. In conclusion, the proposed summarization algorithm can select the more important subspaces, thus extract the most significant key frames to better represent content of the input video.

5 Conclusion

In this paper, a video summarization algorithm using low-rank subspace segmentation is proposed. Using low-rank subspace segmentation, the entire video frames are first grouped into a series of low-rank subspaces, each of which contains frames of similar visual content. Then, by representing the significant subspace with its middle frame, a set of key frames is obtained which then summarizes the given sequence. Experimental results on videos from Open Video Project and comparisons to Open Video storyboard and other existing conventional approaches demonstrate the effectiveness and robustness of the proposed approach.

References

1. Dale K, Shechtman E, Avidan S (2012) Multi-video browsing and summarization. In: IEEE computer society conference on computer vision and pattern recognition workshops (CVPRW), Providence, RI, pp 1–8
2. Avila SEFd, Lopes APB et al. (2011) VSUMM: a mechanism designed to produce static video summaries and a novel evaluation method. Pattern Recogn Lett 32(1):56–68
3. Kim HH, Kim YH (2010) Toward a conceptual framework of key-frame extraction and storyboard display for video summarization. J Am Soc Inf Sci Technol 61(5):927–939
4. Money AG, Agius H (2008) Video summarisation: a conceptual framework and survey of the state of the art. J Visual Commun Image Represent 12(2):121–143
5. Ejaz N, Tariq TB, Baik SW (2012) Adaptive key frame extraction for video summarization using an aggregation mechanism. J Vis Commun Image Represent 23(7):1031–1040
6. Candès EJ, Recht B (2009) Exact matrix completion via convex optimization. Foundations Comput Math 9(6):717–772
7. Keshavan RH, Montanari A, Oh S (2010) Matrix completion from noisy entries. J Mach Learn Res 11:2057–2078
8. Liu G, Lin Z (2010) Robust subspace segmentation by low-rank representation. In: Proceedings of the 27th international conference on machine learning, Haifa, Israel
9. Elhamifar E, Vidal R (2009) Sparse subspace clustering. IEEE Conference on Computer Vision and Pattern Recognition, Miami, pp 2790–2797
10. Kumar M, Loui AC (2011) Key frame extraction from consumer videos using sparse representation. In: 18th IEEE international conference on image processing, pp 2437–2440
11. Pong TK, Tseng P et al (2010) Trace norm regularization: reformulations, algorithms, and multi-task learning. SIAM J Optim 20(6):3465–3489
12. Shi J, Malik J (2000) Normalized cuts and image segmentation. IEEE Trans Pattern Anal Mach Intell 22(8):888–905
13. http://www.open-video.org/
14. Luo J, Papin C, Costello K (2009) Towards extracting semantically meaningful key frames from personal video clips: from humans to computers. IEEE Trans Circuits Syst Video Technol 19(2):289–301
15. Marchionini G, Geisler G (2002) The open video digital library. D-Lib Mag 8(12):1082–9873

Nanostructure Thin-Film Removal via a Cylinders Tool for Computer Touch Sensing Material

P. S. Pa

Abstract This study describes a precise method for the electro-chemical removal of defective tin-doped indium oxide (ITO) thin film from the surface of polyethylene terephthalate (PET) touch screen material using a newly designed multi-cylinder electrode tool. In the current experiment, a small gap-width between the cylindrical cathodes and anodes and the workpiece reduces the time taken to remove a particular amount of ITO by dissolution. A large diameter cathode rotational circle also reduces the time taken for effective ITO film removal because the dissolution effect is facilitated by the provision of sufficient electrochemical power. High rotational speed of the multi-cylinder tool and a high electrolyte flow rate also corresponds with an increase in the dissolution rate of the defective ITO layer. A small diameter of either the anode or cathode combined with enough electric power also results in faster dissolution. The removal rate of the ITO layer is also improved by a reduction in the number of cylinders used. This tool design was found to be most effective for the electrochemical removal of defective ITO thin film nanostructures and can remove the material easily and cleanly in a very short time. This newly designed tool and ultra-precise reclamation process can be used very effectively in the optoelectronics semiconductor industry for the removal of defective ITO layers from PET touch screen substrate to allow the material to be returned to the production line. This recycling will reduce both production costs and pollution.

Keywords Electrochemical · Touch-screen · Multi-cylinder · Electrode · ITO

P. S. Pa (✉)
Department of Digital Content Design, Graduate School of Toy and Game Design,
National Taipei University of Education, NO. 134, Sec. 2, Heping E. Rd,
Taipei City 106, Taiwan
e-mail: myhow@seed.net.tw

1 Introduction

The basic structure of a touch-panel is similar to that of an LCD panel with the addition of two coatings of a transparent tin-doped indium oxide conductive film. The sheet resistance of the $In_2O_3SnO_2$ (ITO) film used on the touch-panel is higher than that of the LCD panel [1, 2]. The primary cause of a decrease in yield rate in LCD production is "dust". When dust particles become attached to the LCD substrate they impair its function and can cause breaks in the circuit, short-circuits, or poor performance. When all the causes of defects in these screens was examined, a large number were found to be dust related [3]. The color filters are the most critical component in LCDs, because each TFT array is matched to a color filter of the same size. This means that the quality of the color filter has a decisive effect on the color reproduction of the LCD. The future of display technology is clearly in flat panel monitors, in which TFT-LCD plays a vitally important role [4].

In the recycling process described here the main difficulty lies in the design of the electrode tool, considering the complicated process of electrochemical metal removal [8]. In addition, workpiece machining through the electrochemical machining (ECM) process can improve precision by the appropriate control of machining conditions or the electrode geometry [5]. The experimental results of Mileham et al. [6] showed that the main factors influencing the quality of an electrochemically machined surface include current density, the flow rate of the electrolyte and the gap width. Brass and zinc alloys can be electro-polished very well in a surface finishing process using PO4-3-P as the electrolyte [7]. Data showed that the gap width between the electrode and workpiece directly influences the current conditions and the discharge of electrolyte dregs [8]. Schuster et al. [9] showed that the machining resolution can be reduced to a few micrometers by applying current pulses of nanosecond duration, and micro structures can easily be machined by ECM. Electrodes of cylindrical shape were developed for use in this recycling system project as very good results have been obtained in the past by careful design of experimental conditions [10, 11].

The cost of materials for most LCD panels of different sizes exceeds 50 % of the total cost of the panels. To reduce material cost, as well as waste and pollution, the implementation of an effective and precise recycling process is one of the hottest issues in the LCD semiconductor industry [12]. A new process has been devised using micro electrochemical removal in a precision etching process to remove defective ITO nanostructure from the surface of sensing material. Using rapidly rotating cylindrical electrodes the process removes the defective layers from the surface of the polyethylene terephthalate digital paper quickly and efficiently. The development of this process was based on both technological and economical considerations. The new precision process is low in cost and highly efficient.

2 Experimental Setup

The material used as the workpiece in these experiments was optical polyethylene terephthalate "digital paper" 400 mm wide and 0.2 mm thick. This PET workpiece, with defective ITO sensing layers was placed in the electrolytic tank and allowed to soak before the recycling process was started, see Fig. 1. The reduction in thickness, after the electrochemical removal of the defective ITO layers from the surface of the optical polyethylene terephthalate substrate, is 20 nm. The experimental setup for the process included an automatic feed device, a DC power supply, a heater, a pump, a flow meter, an electrolytic tank and a filter. The layout of the experiment is shown in the diagram in Fig. 1. The design of the cylinder tool is shown in Fig. 2.

In the current experiment, the electrode arrangement as shown in Fig. 2 was used and continuous DC current was supplied to the anodes to carry out micro
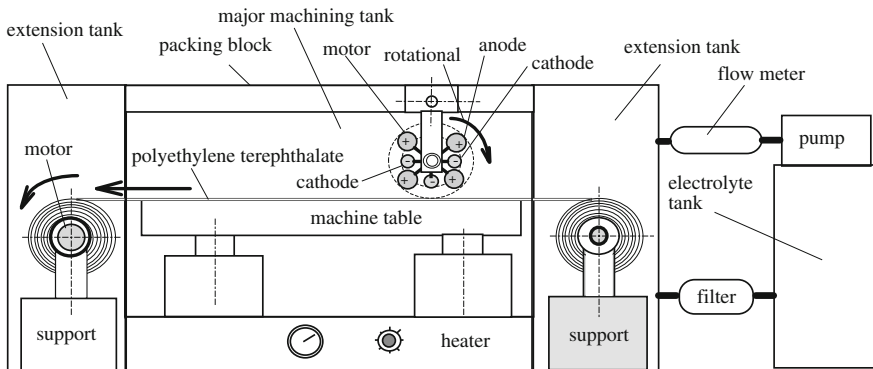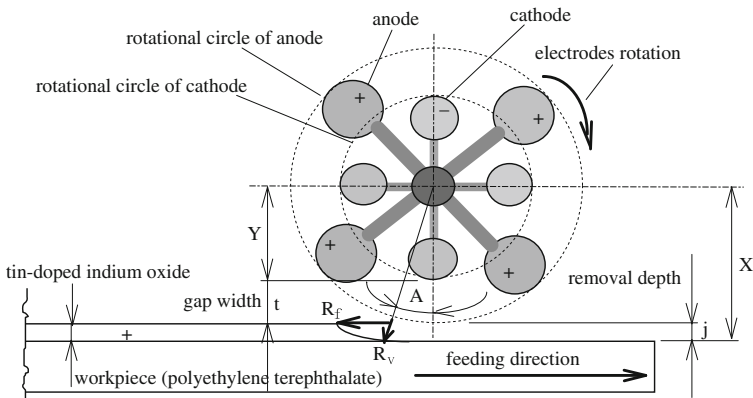

Fig. 1 Experimental setup

**Fig. 2** Configuration of electrodes and workpiece

electroremoval. The electrolyte used was $NaNO_3$ (15 %wt and PO4-3-P 5 %wt). The temperature of the electrolyte was 50 °C. The current rating was 50, 100, 150 and 200 A. The feed rate of the PET diaphragm (workpiece) ranged from 500 to 1,300 mm/min. The diameter of the rotational circle of the cathode was 50, 52, 54 and 56 mm (accompanied by a gap-width between the anode and the cathode was 5, 4, 3, and 2 mm). The diameter of the anodes used was 30, 40, 50 and 60 mm and that of the cathodes was 20, 30, 40 and 50 mm. The cylinders number of the anode was 1, 2, 3 and 4. The cylinders number of the cathode was 1, 2, 3 and 4. The combination of the rotational speed of the cylinder tool and the flow rate of the electrolyte were (200 rpm, 10 L/min), (400 rpm, 20 L/min), (600 rpm, 30 L/min), (800 rpm, 40 L/min). All workpieces were cleaned with water after the reclamation process and dried by air. The amount of ITO film removed was determined by measurements made at more than two locations using a NanoSpec Film Thickness Measurement System (Nanospec Film Analyzer 3000).

3 Results and Discussion

Figure 3 shows that a large cathode rotational circle diameter, a small cathode diameter and a narrow gap between the cathode and the workpiece reduces the amount of time for effective removal of the ITO layers because this is facilitated by an adequate provision of electrochemical power. Examination of the experimental results shows that a small gap accompanied by a large current flow and a fast workpiece feed rate reduces machining time. It can also be seen that a large anode diameter and wider gap increases etching time because the electro removal is limited by a smaller current flow. In this study the most effective and stable
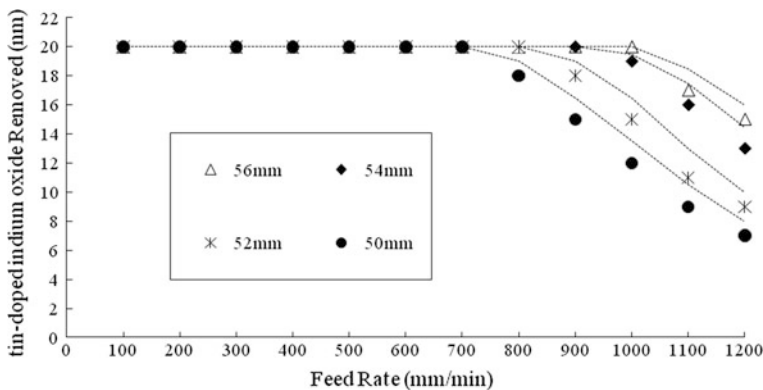
Fig. 3 Amount of micro electro removal at different feed rates of polyethylene terephthalate using different diameter of the ratational circle of the cathode ($NaNO_3$ of 15 %wt and PO4-3 5 %wt, 60 °C, 30 L/min, 150 A, tool 600 rpm)
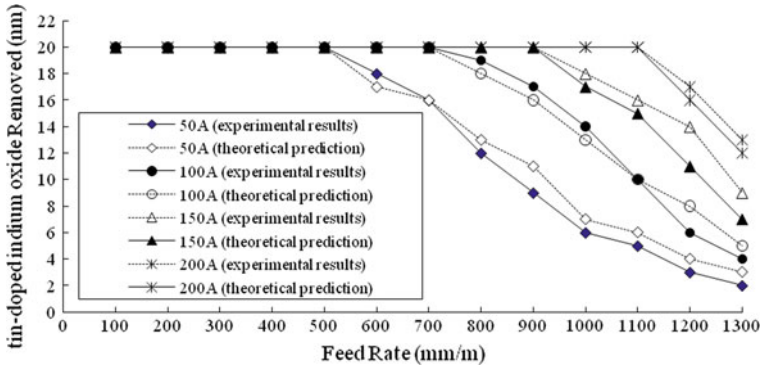
Fig. 4 Amount of micro electro removal at different feed rates of PET using different current rating (NaNO$_3$ of 15 %wt and PO4-3 5 %wt, 60 °C, 30 L/min, tool 600 rpm)

micro electro removal and dregs discharge was achieved with an anode of 20 mm and a gap of 2 mm.

Figure 4 shows that an adequate feed rate of the workpiece combined with enough electric power will completely remove the ITO nanostructure. There is an optimal constant feed rate for the removal of 20 nm of ITO layer from the PET diaphragm that depends on other factors. A fast feed reduces the power delivered to unit area of the plane surface, while a slow feed increases this. In the first case the supply of electrochemical power is insufficient and in the second case the removal time will be increased as will the cost. In order to achieve complete etching of 20 nm of ITO layer the following combination of parameter values is suggested: 50 A at 500 mm/min, 100 A at 700 mm/min, 150 A at 900 mm/min, and 200 A at 1100 mm/min.

According to the formula for the theoretical removal rate on an alloy from Faraday's Law [7],

$$\lambda = \frac{\eta I}{FA\rho(\frac{n_A}{M_A}a_A + \frac{n_B}{M_B}a_B + \cdots)} \quad (1)$$

where η is the efficiency of current, I is the current, t is time, F is the Faraday constant, n_i is the atom number, a_i is the proportion of composition, and M_i is the atomic mass.

Let $\zeta = \lambda/At$, $I_A = \frac{I}{A}$, and $\delta = \zeta/\rho$

$$\delta = \frac{\eta I_A}{F\rho(\frac{n_A}{M_A}a_A + \frac{n_B}{M_B}a_B + \cdots)} \quad (2)$$

$$= \frac{\eta V\sigma}{FA\rho(\frac{n_A}{M_A}a_A + \frac{n_B}{M_B}a_B + \cdots)} \quad (3)$$

where F is the Faraday constant, n_i is the atomic mass, a_i is the proportion of composition, and M_i is the atomic valency, V is the voltage at the gap and σ is the reciprocal resistance, A is the electrochemical machining area, ρ is the density of the workpiece, and R_v is the removal rate in the longitudinal direction. From the above, the theoretical feed rate of the workpiece (PET diaphragm) for the same material removal rate can be calculated.

From Fig. 2:

$$X = Y + t + j \tag{4}$$

where t is the gap width between the cylinder cathode and surface of tin-doped indium oxide (anode); and j is the micro-electro removal depth.

$$\cos A = \frac{t - j}{g} = \frac{Y + j}{Y + t + j} \tag{5}$$

$$(R_f)\sin A = R_v \tag{6}$$

Squaring and simplifying Eqs. (5) and (6):

$$j = \frac{(Y + t)R_v^2}{2(R_f^2 - R_v^2)} \tag{7}$$

where R_f is the feed velocity of the workpiece and R_v is the removal rate in the longitudinal direction. From Eq. (7):

$$j = \frac{(Y + t)\left[\frac{\eta I}{(FA\rho)^2(\frac{n_A}{M_A}a_A + \frac{n_B}{M_B}a_B + \cdots)^2}\right]}{2R_F^2 - 2\left[\frac{1}{(FA\rho)^2(\frac{n_A}{M_A}a_A + \frac{n_B}{M_B}a_B + \cdots)^2}\right]^2} \tag{8}$$

where A is the electrochemical machining area, I is current rating, ρ is the density of the workpiece. From the above, the theoretical feed rate of the workpiece (PET diaphragm) for the same material removal rate can be calculated, where η, I, F, and A are regarded as constant for the material. From Eq. (8), the experimental results were found to agree with the theoretical prediction (see Fig. 4). The experimental results show the removal depth (j) of the ITO layer is directly proportional to the current rating (I) and in inverse ratio to the feed rate of the workpiece (R_f), this is also in agreement with the theoretical prediction (see Fig. 4).

Figure 5 illustrates that the performance of a small diameter anode with a small diameter cathode is facilitated by the provision of sufficient electrochemical power. From Eqs. (2) and (8), the feed rate of the workpiece (δ) is directly proportional to the removal depth of 20 nm, the removal depth (j) is directly proportional to the current rating (I), and the current rating (I) is in inverse proportion to the electrochemical machining area (A). A small diameter anode, or a small diameter cathode, working over a small electrochemical machining area (A)
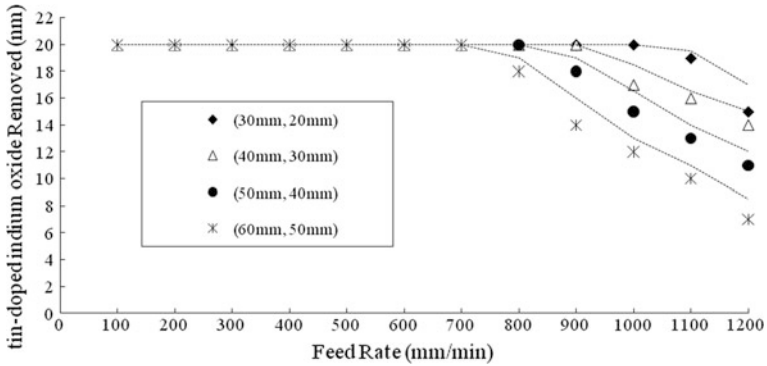
Fig. 5 Amount of micro electro removal at different feed rates of polyethylene terephthalate using different diameter of the anode and cathode ($NaNO_3$ of 15 %wt and PO4-3 5 %wt, 60 °C, 30 L/min, 150 A, tool 600 rpm)
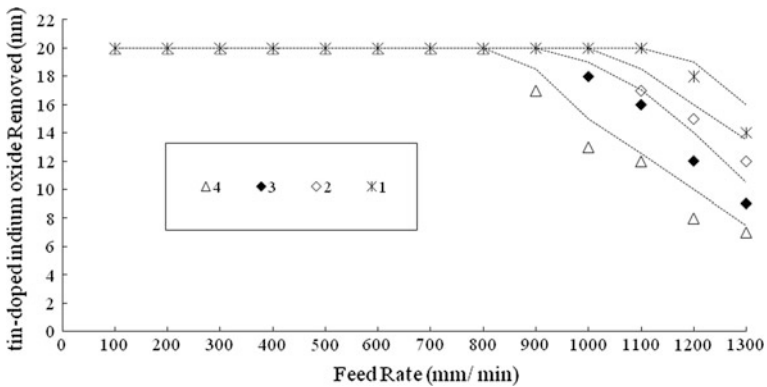
Fig. 6 Amount of micro electroremoval at different feed rates of polyethylene terephthalate using different pins number of the anode and cathode ($NaNO_3$ of 15 %wt and PO4-3 5 %wt, 60 °C, 30 L/min, 150 A, tool 600 rpm)

with high current density ($\frac{I}{A}$) and a fast workpiece feed rate (δ), will reduce the machining time. Figure 6 show that the removal rate of the ITO thin film is improved by a decrease in the number of cylinders to increase the electrochemical power density on the machining area. The removal of dregs from the machining gap is also more efficient as a result of the increase in the amount of open space that also allows heat to be removed more efficiently. The use of a smaller number of anodes and cathodes with a faster.

4 Conclusion

The precision recovery of touch screen polyethylene terephthalate (PET) substrate by the micro electro removal of defective ITO nanostructure, and the design of a multi-cylinder electrode tool in the current study is of major interest to the LCD optoelectronics industry. In this process defective products can be reclaimed and returned to the production line with considerable saving of waste, pollution and cost. The experimental work has shown that a high current flow and fast workpiece feed effectively speeds up the removal effect. A large cathode rotational circle diameter accompanied by a small gap between the cathode and the workpiece reduces the time taken to remove a fixed amount of ITO layer. The removal rate of ITO thin film is also improved by a decrease in the number of cylinders. A small diameter anode, or a small diameter cathode, a high electrolyte flow rate combined with a higher rotational speed of the electrode tool all correspond to a more rapid removal of ITO nanostructure.

References

1. Lai CC, Tsai CC (2004) Neural calibration and Kalman filter position estimation for touch panels. In: Proceedings of the 2004 IEEE international conference on control applications, vol 2, pp 1491–1496
2. Tsai DM, Lin PC, Lu CJ (2006) An independent component analysis-based filter design for defect detection in low-contrast surface images. Pattern Recogn 39(9):1679–1694
3. Daeil K, Steven K (2002) Effect of secondary ion beam energy and oxygen partial pressure on the structural, morphological and optical properties of ITO films prepared by DMIBD technique. Int'l J Surf Coat Technol 154(2–3):204–208
4. Iwase H, Murata A (2002) Empirical study on improvement of usability for touch-panel for elderly-comparison of usability between touch-panel and mouse. In: 2002 IEEE international conference on systems, man and cybernetics, vol 1, pp 252–257
5. Mc Geough JA (1974) Principles of electrochemical machining. Chapman and Hall, London, pp 1–10
6. Mileham AR, Harrey SJ, Stout KJ (1986) The characterization of electrochemically machined surfaces. Wear 109:207–214
7. Iskander SS, Mansour IAS, Sedahmed GH, Fawzy WA (1980) Electropolishing of brass alloys in phosphoric acid. Surf Technol 10(5):357–361
8. Datta M, Landolt D (1981) Electrochemical machining under pulsed current conditions. Elector Acta 26:899–907
9. Cagnon L, Kirchner V, Kock M, Schuster R, Ertl G, Gmelin WT, Kuck H (2003) Electrochemical miromachining of stainless steel by ultra short voltage pulses. Z Phys Chem 217:299–313

10. Pa PS (2008) Precision removal of ITO layer using plate-form tool design. J Solid State Electrochem 12(11):1445–1451
11. Kim BH, Ryu SH, Choi DK, Chu CN (2005) Micro electrochemical milling. J Micromech Microeng 15:124–129
12. Daeil K, Steven K (2002) Effect of secondary ion beam energy and oxygen partial pressure on the structural, morphological and optical properties of ITO films prepared by DMIBD technique. Int'l J Surf Coat Technol 154:204–208

Research on Cow Epidemic Monitor System Based on ArcGIS Engine

Wenbo Wang, Hongli Zhou and Jinghong Li

Abstract In order to improve the work environment for cows epidemic monitor and raise working efficiency, this paper propose a general design scheme for this system and develops partial function using the methods of function analysis. Information system analysis design,with the VS.NET,ComGIS development platform(ArcGIS Engine),and database technology, Using various method of spatial analysis, realized the comprehensive analysis of the historical cows disease and Space–time distribution having certain effect for improving the level of controlling cows disease, providing some reference for cows disease geographic information system construction of relevant departments.

Keywords GIS · AE · C# Language · Cows Epidemic

1 Introduction

At present, the cows epidemic monitor method of china that breeding point found animal epidemic report to the superior department, Animal medical staff of superior department to breeding site to check statistics, formulate relevant prevention and control measures. Traditional way lack of efficient and scientific on

W. Wang (✉) · H. Zhou
Anhui University of Science and Technology, Anhui 232001 Huainan, China
e-mail: 11250456@qq.com

J. Li
Henan Electric Power Corporation Maintenance Company, Henan 450000 Zhengzhou, China

control epidemic, Using GIS technology, computer technology and database technology to cows epidemic monitor system research is relatively less. Therefore, this article takes the heilongjiang province as an example, to discusses the use of GIS (ArcGIS Engine), computer and database technology construction cow epidemic monitor system, to explore the new ideas and new methods of the cows epidemic monitor in order to solve the low efficiency problem of routine work, aimed at reduce the cows epidemic monitoring work labor intensity, improving information management of epidemic, standardization and efficience, providing efficient and quick decision basis for related epidemic management departments.

2 ArcGIS Engine Summarize

You can choose a different environment develop ArcGIS Engine (AE) application program, such as Microsoft Visual Studio,Delphi,ECLIPSE,Sun One Studio and so on.

Firstly, registered AE development components in the integrated development environment, then build a application based on form, add AE components, choice a model constructing our own application. The main technical points: (1) Form design and layer import: In C# environment loading controls such as MapControl, PageLayout, TOCControl, ToolbarControl and related references. Add controls on form, MapControl, PageLayout controls to realize function of Introducing layer, get relevant interface which provided by component library, so as to realize all kinds of function.(2) Code control: Set TOCControl, ToolbarControl and PageLayout control connection to realize the function, Through the code control, call AE developer kit and execute procedures that can obtained small application system [1].

2.1 Simple Introduction of ArcGIS Engine

ArcGIS Engine(AE) is used to create custom application of a complete embedded GIS component library, AE contains two parts:

ArcGIS Engine Developer kit–developers use it to construct customizing application;

ArcGIS Engine Runtime–final customer can run application program in their computer which including ArcGIS Engine component; AE Developer kit is an independent development product, In fact, AE is composed of a set of core ArcObjects kit, can be called in various programming interface l. AE oriented to the program developer of GIS project not final customer, AE is not a terminal application for developers just component suite with secondary development function for the development of new application program.

AE Runtime deployment requirements each software seats need separate operation environment license, In addition to the basic function of the options, there arefour runtime options: spatial runtime option, 3D runtime option, geodatabase update runtime option and network analysis option [2].

2.2 Development Method of AE

You can choose a different environment develop AE application program, such as Microsoft Visual Studio,Delphi,ECLIPSE,Sun One Studio and so on [3].

Firstly, registered AE development components in the integrated development environment, Then build a application based on form, add AE components, choice a model constructing our own application. The main technical points including:

1. Form design and layer import: In C# environment loading controls such as MapControl, PageLayout, TOCControl, ToolbarControl and related references. Add controls on form, MapControl, PageLayout controls to realize function of Introducing layer, get relevant interface which provided by component library, so as to realize all kinds of function.
2. Code control: Set TOCControl, ToolbarControl and PageLayout control connection to realize the function, through the code control, call AE developer kit and execute procedures that can obtained small application system.

3 Overall Design of System

3.1 Goal of System Design

Goal of System Design (GIS) technology combined with database technology, to analysis related information of municipal (county) monitoring station and veterinary epidemic prevention station of dairy cow disease, to regular and systematic statistics information that cow disease epidemic monitoring and epidemic prevention management and expert material of heilongjiang province, to grasp information dynamic and realize information resources visualization management [4].

Using a variety of spatial analysis methods to realize a variety of functions which include the history of dairy cow disease, comprehensive analysis of the time and space distribution rule, comprehensive analysis of specific regional ecological environment and cause of disease, disease prevention and control and decision-making of reasonable allocation of resources, and so on. Based on the information of the deep utilization and analysis, understanding the disease development law of

cow for government management departments, providing scientific basis for management departments, carry out prevention and control of the cow disease. The system provide real-time service for dairy producers and decision makers, improve the management efficiency and technology promotion level of modernization.

3.2 Integrated Frame of System

Using Microsoft Visual Studio 2010 as a development tool, using C# as development language, database choose SQL Server2008 platform, using AE components as secondary development basic platform. Integrated frame of system as Fig. 1 shown.

3.3 Function Design of System

3.3.1 Function Module Design of System

According to the functional requirements of the cow epidemic monitor system, the whole system can be divided into four modules. Each function modules can achieve the following functions:

1. Graphics management function module: the superiority of the GIS system is mainly embodied in data content and quality, the cow disease related information distribution range is wide, cows point, disease information constantly updated, users can easily entry, modify, epidemic related information and basic
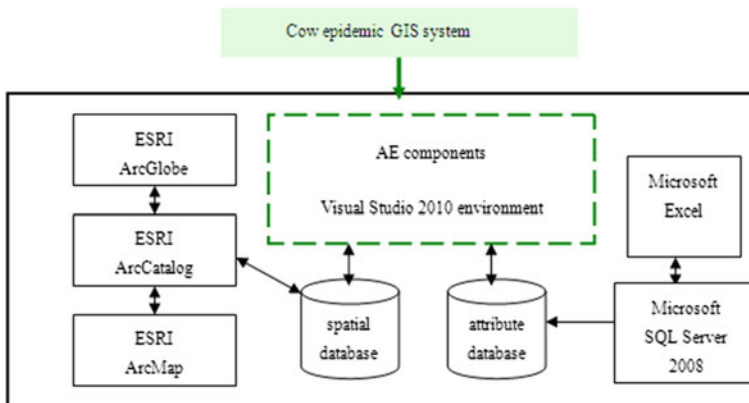

Fig. 1 Integrated frame of system

data of cows point, establish or modify the corresponding space and attribute database.

2. Information query, display function module: final user can specify conditions of selection, designated retrieval area, graphics conditions, attribute conditions; the result of inquires can output by kinds of ways, such as pictures text and table.

3. Information statistic analysis function module: Statistical analysis including simple statistics and comprehensive statistics. Simple statistics include the attribute data of mean value, maximum and minimum value calculation; comprehensive statistical realize histogram, point density diagram, pie chart and graph generation, intuitive statistical results can be seen by users.

4. Disease emergency auxiliary support function module: cow disease Combined with space and time information, the spatial data and attribute data such as cow epidemic space position, space distribution and epidemic dynamic, which can be used to overlay analysis, density analysis and surrounding condition analysis and dynamic analysis. So as to form cow disease status report, cow varieties area evaluation, then for milk cow industry relevant departments to provide scientific basis for cow epidemic emergency strategy formulation, dairy industry layout planning.

3.4 Database Design of System

Database of cow epidemic GIS system including: the space database and attribute database, respectively store spatial information and attribute information of the cow epidemic.

3.4.1 Construction of Spatial Database

Spatial database is the core of the whole system, the main storage research regional spatial data. Spatial data use GeoDatabase to management in this system.

3.4.2 Construction of Attribute Database

Attribute data is also called the statistical data or project data. It is the detailed description of attributes exclusive spatial feature, it contains the description of target type, the specific instructions and description. Most of GIS software directly records spatial data, in order to symbolic and layered display.

This system uses the small desktop database management system-SQLsever2008, which relatively simple structure, data revised and update more conveniently, maintenance and usage easily [5].

4 System Function Realization

This paper using ArcGIS as map modification editor software platform, using C# as development language, using GeoDatabase for spatial data management, realize system function development, software simulation run results is good.

4.1 Main Interface of System

System interface is human–computer interaction interface, including how the users command system and system how to submit information to the user. Friendly user interface allows users to grasp the system more easily, increase acceptance the system level of, the system main interface as Fig. 2 shown.

4.2 Example Introduce of System Function

The core function of the system is to dynamic monitoring cow epidemic, epidemic situation data of time and space statistical analysis, form thematic map according to the relevant data. Provide decision basis for related sections.
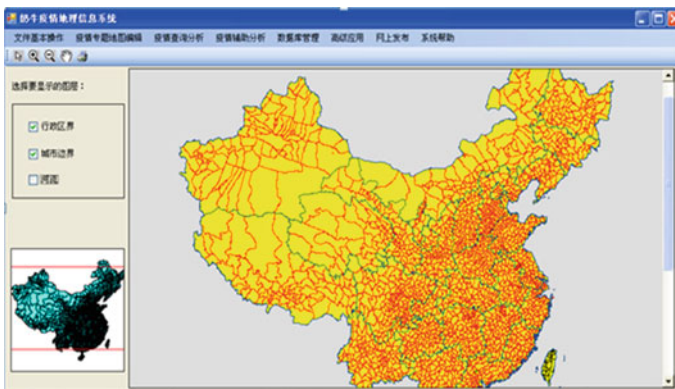
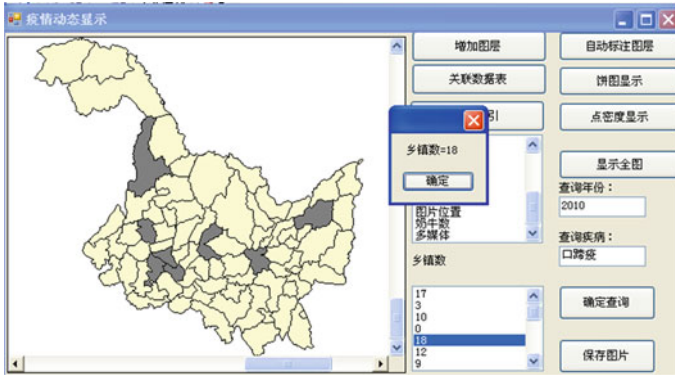

Fig. 2 Main interface of system

Fig. 3 Dynamic display query of cow diseases information

4.2.1 Query Information of Cow Disease

Cow disease related information dynamic display that is epidemic dynamic changes for a period of time, analyzes the developing trend of history. Through the Fig. 3 shows that change different interface button to get what you want inquires.

4.2.2 Design of the Thematic Map

Thematic map which use special image symbol system and unique representation method, according to the specific map object and the service target, outstanding display one or several spatial elements. The system generated thematic map for the cows relevant departments to provide visual information and epidemic dynamic data, provide support for the department to make scientific and reasonable prevention and control plans Fig. 4.
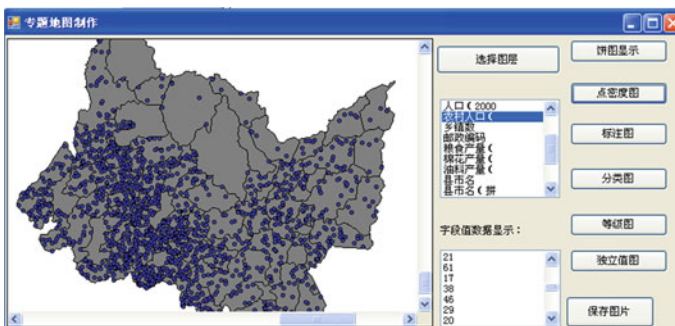

Fig. 4 Dot density map

5 Summary

This paper has deeply researched cows epidemic monitor system, using.net sharing Microsoft Visual Studio 2010 platform and GIS product AE components developed software system. The work demand of epidemic monitoring has included in the system, related information of relational graph data and attribute data has realized digitization; the cost of the epidemic situation data update work can be reduced and efficiency can be improved. System with flexible and simple operation, and can provide comprehensive detailed information of animal epidemic to management department, and also can provide reliable basis for government departments make decisions.

However, this system still need further perfect in some ways, the data management function still need to further research, the system stability still need to further improve.

References

1. Gang L, Xiaodong L et al. (2004) The geographic information system secondary development examples tutorial. The first edition. Tsinghua university press, Beijing, pp 91–92
2. Gold CM, Edwards G (1992) The voronoi spatial data model: two and 3d application in image analysis. ITC Journal, pp 11–19
3. Microsoft Corporation (2002) Microsoft visual C#.NET language reference. Microsoft Press, US
4. Yingchang L (2004) The geographic information system in the field of health of the progress in the application of foreign medical. Medical Geogr Fascicle 04:97–101
5. Jinsong Ma (2008) Introduction to geographic information system. The first edition. Higher education press, Beijing, pp 12–23

Application of Dual Bilinear Interpolation Fuzzy Algorithm in Fan Speed Control

Kai-feng Huang, Ze-gong Liu, Jing Yang, Feng Xu, Kui Gao and Ya Kang

Abstract This paper focused on ordinary fuzzy controller for operation of control rule by approximate processing for a particular discrete points, resulting in steady-state performance results information loss; made binary optimized fuzzy control algorithm of bilinear interpolation. On the basis of basic fuzzy controller, it was to avoid error due to an approximate calculation of the loss of information, used by dual bilinear interpolation algorithm to improve continuity of control rules, so as to achieve the goal of improving stability precision.

Keywords Fuzzy control · Dual bilinear interpolation · Fan

1 Introduction

Fuzzy control, as a matter of intelligent control technology, since United Kingdom scholars Mamdani used to the control of boiler and steam engine in 1974, fuzzy control theory and technologies have come a long way from the development. It have been successfully used in high-interference, coupling, time-varying, Multi-variable systems. It has the ability of adaptive control, especially suitable for difficult to establish mathematical models of systems. However, output of ordinary fuzzy controller was a table, its smoothness and accuracy is not guaranteed.

K. Huang (✉) · Z. Liu · K. Gao · Y. Kang
School of Energy and Safety, Anhui University of Science and Technology, Huainan 232001, China
e-mail: kaifenghyj@163.com

K. Huang
Information and Electrical Engineering, Department of Huainan vocational and technical College, Huainan 232001, China

J. Yang · F. Xu
School of Science, Anhui University of Science and Technology, Huainan 232001, China

Professor Li hong-xing first pointed out that fuzzy controller is essentially interpolation in 1997. This assertion, normal fuzzy control algorithm is essentially one kind of interpolation methods, provided a very good way to improve precision of fuzzy controller and rich fuzzy controller design method. Dual bilinear interpolation algorithm, which is based on fuzzy control with smooth and completely fuzzy rule as conventional fuzzy controller, formed mapping from input to output directly, Various fuzzy control algorithm, based on theory of interpolation algorithm, could using it on rules-refinement real-time, to overcome slow response, long adjustment time, to solve membership function graphics adjustment problem, and other shortcomings of the normal fuzzy control, when complete rules library conditions was satisfied. Dual bilinear interpolation algorithm has been larger improved in precision of controller [1, 2].

2 Normal Structure of Two-Dimension Fuzzy Controller

Normal structure of two-dimension fuzzy controller as shown in Fig. 1. Where e and ec (input variable) were errors and bias changes of system respectively, u (output variable) was output of controller.

Two-dimension fuzzy controller design is as follows: universe of discourse of e, ec (input variable) and u(output variable) are {−5,5}; Language is represented as {negative large, negative small, zero, positive small, positive large}. Membership function of Two-dimension fuzzy controller as shown in Figs. 2 and 3. Designed fuzzy rule as shown in Table 1 fuzzy inference system use Mamdani-type algorithm defuzzification use method of centroid, we can get table of the fuzzy control output as shown in Table 2 [3].
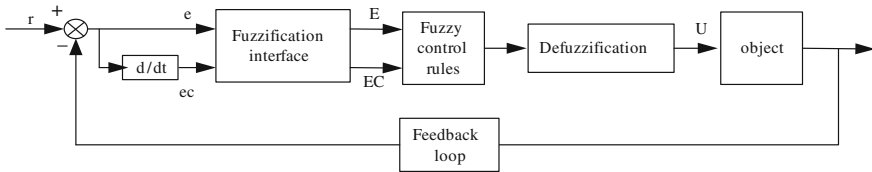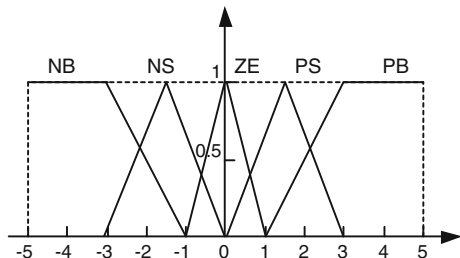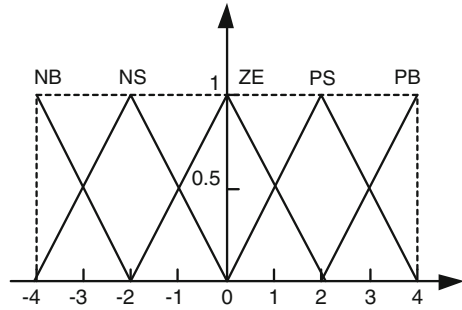

Fig. 1 Two-dimension fuzzy controller

Fig. 2 Membership function of E, EC

Fig. 3 Membership function
of U

Table 1 Fuzzy rule

E	EC				
	PB	PS	ZE	NS	NB
PB	PB	PB	PS	PS	ZE
PS	PB	PS	ZE	ZE	NS
ZE	PS	PS	ZE	NS	NS
NS	PS	ZE	ZE	NS	NB
NB	ZE	NS	NS	NB	NB

Table 2 Table of output of the fuzzy control

E	EC										
	−5	−4	−3	−2	−1	0	1	2	3	4	5
−5	−4	−4	−4	−3.5	−3.5	−2	−2	−1	0	0	0
−4	−4	−4	−4	−3.5	−3.5	−2	−2	−1	0	0	0
−3	−4	−4	−4	−3.5	−3.5	−2	−2	−1	0	0	0
−2	−3.5	−3.5	−3.5	−2.2	−2.2	−2	−1	0	1	1	1
−1	−3.5	−3.5	−3.5	−2.2	−2	−2	0	1	2	2	2
0	−2	−2	−2	−1	0	0	0	1	2	2	2
1	−2	−2	−2	−1	0	0	2	2.2	3.5	3.5	3.5
2	−1	−1	−1	0	1	1	2.2	2.2	3.5	3.5	3.5
3	0	0	0	1	2	2	3.5	3.5	4	4	4
4	0	0	0	1	2	2	3.5	3.5	4	4	4
5	0	0	0	1	2	2	3.5	3.5	4	4	4

3 Dual Bilinear Interpolation Algorithm

Implementation process of dual bilinear interpolation algorithm optimize fuzzy control algorithm [4].

Step 1: e = x, ec = y, $\forall(X, Y)$, we can get $\tilde{X}'$, $\tilde{Y}'$ from fuzzification, and then we can get $\widetilde{X}_c$, $\widetilde{Y}_c$ from $<\tilde{X}'>$, $<\tilde{Y}'>$ for rounding operation. At the same time as structures are:

$A1\left(X_c\widetilde{}-1,\ Y_c\widetilde{}-1\right),\ A2\left(X_c\widetilde{}-1,\ Y_c\widetilde{+}1\right),\ A3\left(X_c\widetilde{+}1,\ Y_c\widetilde{+}1\right),\ A4\left(X_c\widetilde{+}1,\ Y_c\widetilde{}-1\right),$

Step 2: Select $\widetilde{X}_c$, $\widetilde{Y}_c$ enter rule query tables, we can get $\widetilde{Z}$, and we can get $\widetilde{Z}_1$, $\widetilde{Z}_2$, $\widetilde{Z}_3$, $\widetilde{Z}_4$ on same way.

Step 3: Dual bilinear interpolation. There are four factors in Function of Dual bilinear interpolation. So we must select four point in rectangle element, and then we can learn that function value of four point as condition of interpolation, Rectangle unit as shown in Fig. 4. We will select four point in rectangle element:

$A1\left(X_c\widetilde{}-1,\ Y_c\widetilde{}-1\right),\ A2\left(X_c\widetilde{}-1,\ Y_c\widetilde{+}1\right),\ A3\left(X_c\widetilde{+}1,\ Y_c\widetilde{+}1\right),\ A4\left(X_c\widetilde{+}1,\ Y_c\widetilde{}-1\right),$

Enter rule query tables, we can get $\widetilde{Z}(\widetilde{X},\ \widetilde{Y})$. Rectangular element anywhere through coordinate transformation into the centroid c-the origin of the standard unit e, coordinate transform and inverse transform, respectively.

$$\begin{cases} \xi = \frac{\widetilde{X}-\widetilde{X}_c}{a} \\ \eta = \frac{\widetilde{Y}-\widetilde{Y}_c}{b} \end{cases};\ \begin{cases} \widetilde{X} = \widetilde{X}_c + a\xi \\ \widetilde{Y} = \widetilde{Y}_c + a\eta \end{cases}$$

Construct interpolation function based on the natural coordinate system functions, $N_1(\xi,\eta)$, $N_2(\xi,\eta)$, $N_3(\xi,\eta)$, $N_4(\xi,\eta)$ correspond respectively to the points A_1, A_2, A_3, A_4.

Dual bilinear interpolation basis function in rectangle element is: $P(\xi,\eta) = N_1(\xi,\eta)Z_1 + N_2(\xi,\eta)Z_2 + N_3(\xi,\eta)Z_3 + N_4(\xi,\eta)Z_4$ $\widetilde{Z}_1$, $\widetilde{Z}_2$, $\widetilde{Z}_3$, $\widetilde{Z}4$ are   output value   of $A1\left(X_c\widetilde{}-1,Y_c\widetilde{}-1\right),\ A2\left(X_c\widetilde{}-1,Y_c\widetilde{+}1\right),\ A3\left(X_c\widetilde{+}1,Y_c\widetilde{+}1\right),$ $A4\left(X_c\widetilde{+}1,Y_c\widetilde{}-1\right)$. By table look-up, dual bilinear interpolation basis function $N_i(\xi,\eta)$ must satisfy this condition:

(1) $N_i(\xi,\eta)$ is dual bilinear function about ξ and η;

Fig. 4 Rectangle element
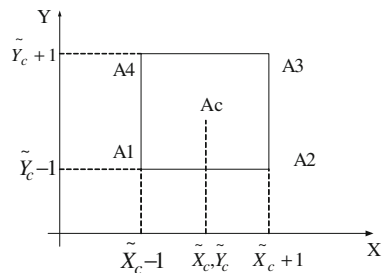
(2) $N_i(\xi_j, \eta_j) = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$ $(i, j = 1, 2, 3, 4)$, (ξ_j, η_j) is point in natural coordinates.

By the above condition, assume

$$N_1(\xi, \eta) = c(X_c \tilde{-} 1 - \xi,) (Y_c \tilde{-} 1 - \eta),$$

Among c is undetermined factor. We can get c into the formula

$$N_1(X_c \tilde{-} 1, \ Y_c \tilde{-} 1) = Z_1$$

Then

$$\begin{cases} N_1(\xi, \eta) = c(X_c \tilde{-} 1 - \xi,)(Y_c \tilde{-} 1 - \eta) \\ N_2(\xi, \eta) + c(X_c \tilde{-} 1 + \xi,)(Y_c \tilde{+} 1 - \eta) \\ N_3(\xi, \eta) = c(X_c \tilde{+} 1 + \xi,)(Y_c \tilde{+} 1 + \eta) \\ N_4(\xi, \eta) = c(X_c \tilde{+} 1 - \xi,)(Y_c \tilde{-} 1 + \eta) \end{cases}$$

Then dual bilinear interpolation function in standard element is:

$$P(\xi, \eta) = \sum_{i=1}^{4} c(1 + \xi\xi_i)(1 + \eta\eta_i)Z_i$$

4 Simulation Results

Main method of the ordinary fuzzy controller execute control ruler is the off-line table look-up. The off-line table look-up method only carries on the table look-up to some specific points, the input and the output all is the specific dispersed points, other points are approximate through processing for these specific separate spots. The result can create the partial information to be lost. Therefore, the stable performance is not good. The dual bilinear interpolation optimizes the fuzzy control algorithm is one kind of continuous fuzzy controller. In the basic fuzzy controller foundation, it avoids the error on account of the approximate calculation lose the information, with through dual bilinear interpolation method enhancement control rule continuity, thus achieved increases the stable state precision the goal. This kind of control algorithm common fuzzy controller has following 4 merits:

1. It can carry out the control rule correctly.
2. It may reduce the static error.
3. Reflection information change sensitivity is high.
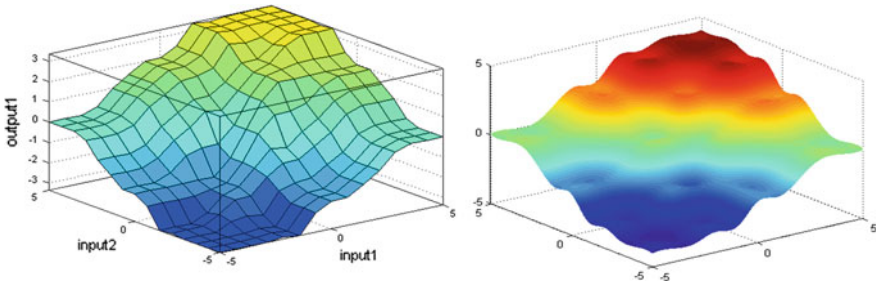4. Precisions are higher, the smooth degree is better, conforms to human's thought.
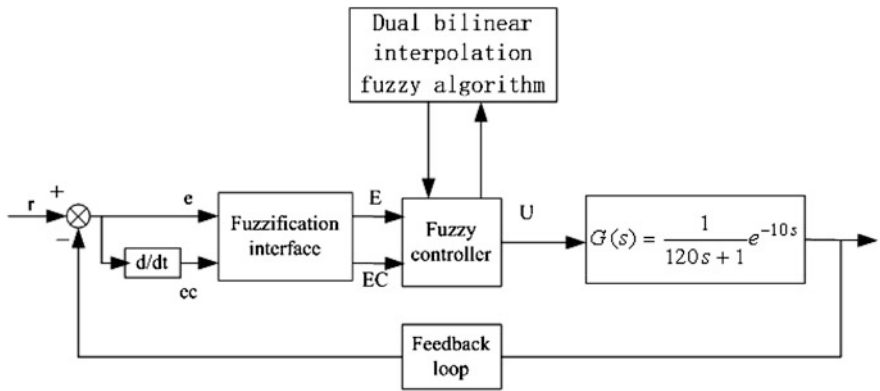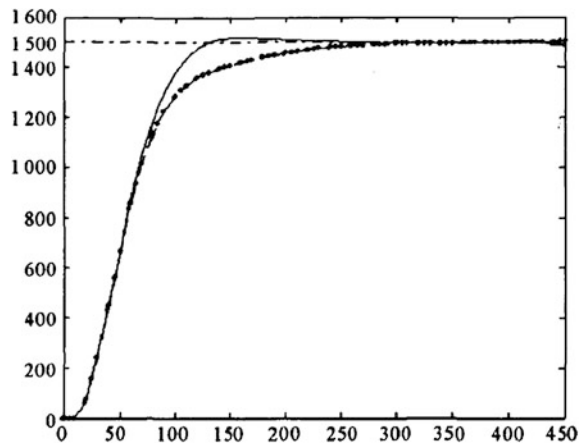
Fig. 5 Comparison of output

Fig. 6 Structure of dual bilinear interpolation fuzzy system

Fig. 7 Response curve

5 Application of Dual Bilinear Interpolation Fuzzy Algorithm in Fan Speed Control

For verifying practicability of Dual bilinear interpolation fuzzy algorithm, we will apply this theory to fan speed Simulation control as shown in Figs. 5 and 6.

The mathematics model of fan speed control is $G(s) = \frac{1}{120s+1} e^{-10s}$ [5–8], when demand of wind pressure is 1,500 Pa, Simulation curve as shown in Fig. 7.

References

1. Huang K, Yangjing, Li Z (2006) A fuzzy controller for coalmine gas mobile pumping-exhaust system. J Huainan Vocat Tech Coll 7(22):7–10
2. Huang K, Yang Y (2010) A controller for two coalmine gas mobile pumping-exhaust system based on embedded system. J Huainan Vocat Tech Coll 10(35):39–42
3. Zhao JX, Jin HZ (2010) Gas sensor fault diagnosis based on wavelet packet and neural network. Transducer Microsyst Technol 29(5):80–82
4. Yang WG, Zhao HL (2009) Fuzzy control algorithm based on spline interpolation. Fuzzy Syst Math 23(3):152–157
5. Chen X, Li XW, Fang XP, Liu B (2009) Design and simulation of fuzzy predictive controller in fan speed regulation system. J Shandong Univ Sci Technol Nat Sci 28(5):49–53
6. Chen G (2006) Optimizing of support vector machine time series forecasting model parameters based on genetic algorithms. Chin J Sci Instrum 27(9):1180–1184
7. Goldberg D (1989) Genetic algorithms in search, optimization and machine learning. Addison-Wesley, Reading, MA
8. Zhou GB, Zhu ZC, Chen GZ (2010) Fault diagnosis of gas sensor based on wireless sensor network. J Vibr Meas Diagn 30(1):23–27

Research on Credibility of Web Service Composition Based on Stochastic Petri Net

Juan Guo, Xianwen Fang, Dawei Zheng and Xiuyun Zhang

Abstract The service-oriented architecture (SOA) has become inevitable in the development and wide application of computers and networks. SOA is characteristic of service composition so that it can provide more services to companies. While there is a promising future for the application of web service composition, some issues such as safety and reliability hinder its further development. To increase the trustworthiness of web service composition, the paper puts forth some indexes and a Petri net based approach. The credibility of web service composition can be evaluated with this approach.

Keywords Credibility · Stochastic Petri Net · Web service composition · Markov chain

1 Introduction

The web service, an application program, is a kind of distributed and autonomous modularization. And by combining single web services in the network it can provide more complex value-added services. However, web service composition also suffers from some defects due to some uncertain factors of the network. In recent years, web service composition and its evaluation have become a researching focus for many scholars. In the paper, the writer has a thorough review of literatures concerned. Ref. [1] classify different composition methods of web service. Ref. [2] studies the finding and composition of dynamic web service on the basis of semantic web service. Ref. [3] suggests methods on how to bind web service composition through safety performance. Refs. [4–6] apply the Petri net in

J. Guo (✉) · X. Fang · D. Zheng · X. Zhang
Anhui University of Science and Technology, Anhui 232001 Huainan,
People's Republic of China
e-mail: guojuan_aust@foxmail.com

evaluating safety and credibility of web service or service composition. Ref. [7] researches into the judgment of the consistency and compatibility of web service through MTT. Ref. [8] evaluates the trustworthiness of SOA software by a software credible paradigm and an algebra model. The focus of the present paper is the credibility of service composition. Credibility is defined in Refs. [9, 10] as a capacity to avoid frequently invalid services. In Ref. [11], it is defined in another way as the system's ability to continuously provide with services and resume services at any given time when it is invalid due to external factors, human mistakes, environmental influences, or hardware failures. In Ref. [11], it is believed that the performance and result of the reliable network system can be predicted by resolving its model, with its performance monitored, its result evaluated, and exceptions controlled. Based on the definition of credible network systems from Ref. [11], the paper comes up with evaluation indexes for web service composition.

The Petri net is being used by more and more researchers. Besides, the Petri net is suitable for describing the relations like sequence, concurrency, conflict and synchronism of processes and ingredients in the system. Stochastic Petri Net (SPN) takes into account the delays by transition so that it can be used to model and analyze uncertain systems. In this paper, the research on web service composition is conducted through expanding SPN, modeling SPN and getting indexes.

2 Web Service Composition Based on SPN of Information Extension

2.1 SPN of Information Extension

The study on the performance of web service composition is usually done through constructing models of SPN, reachability graphs, and Markov chains. Then some analyzing and computing methods are used to assess the performance of the model. As for how to analyze the performance of SPN, refer to Ref. [12].

For the sake of analyzing the credibility of web service composition, SPN is extended by adding data places and transition-attaching information.

1. Data places.

When doing model analysis, considering model mistakes triggered by data flow, a place type is introduced called data places on the basis of the original SPN. Data places demonstrate how data are transmitted between services. If there is a token, it means that the data are ready, and vice versa. The capacity of the data place is 1, shown by dotted small circles. The firing rules remain the same after the introduction of data places.

2. Transition-attaching information

When analyzing the dependability and security of the model, transitions are expanded with relevant information. λ is a set of transition firing rate, which expresses the average time spent web service components carry out effective services. u is a set of transition failure rate, referring to the average time spent when web service components fail due to attacks; β is a set of transition repair rate, indicating of the average time needed when web service returns to normal from failures.

2.2 Web Service Composition Net

In order to make use of the Petri net to describe the web service composition, it is necessary that the components of web service composition correspond to the Petri net. In the following part, web service composition is defined based on some previous researches.

Definition 1: The web service S can be defined as a seven-tuple: $S = (Id, Name, Desc, URL, Oper, Time, Defeat)$.

The meaning of each one of the seven components is as follows. *Id*: the unique identifying of the web service; *Name*: the name of the web service; *Desc*: description of the web service; *URL*: the call for the web service; *Oper*: operation set of the web service; *Time*: average executive time of web service; *Defeat*: average failure time of web service when facing vicious attacks or other problems.

Definition 2: The web service composition net (WSCN) is defined as an eight-tuple, which is used to model for the dynamic performance of service composition. $WSCN = (S, T; F, \lambda, \beta, u, i, o)$. S means finite sets of data places, reflecting the state of web service. T refers to finite transition set, standing for operations in and between web services. F, is about the relation between service states and operations. λ is the sets of the average time of implementing transition. β is the repair rate after transition failures. u is the rate sets of transition failures. i means the input place, marking the beginning of service composition procedures. o is the output place, showing the ending of service composition procedures.

Definition 3: For the model of web service composition to be correct, it must meet the following basic demands:

1. There is an input place i and an output place o in every model;
2. Any transition is possible to be fired;
3. Under any circumstances, web service composition can be ended;
4. At the ending time, only the output place contains tokens;
5. When the transition is enabled, the input data of it must be usable;
6. There exists no concurrent activities sharing the same output variants;

3 Credibility of Web Service Composition

The paper puts forward three indexes according to the characteristics of web service composition: correctness, dependability and safety.

3.1 Analysis on the Correctness of Web Service Composition Modeling

With regards to the correctness analysis of web service composition, most researches focus on the correctness of composition's flow, but no research estimates the mistakes evoked by data. This paper takes into account errors caused by data in analyzing the correctness of web service composition, mainly including errors of invalid data and data collision which is described in definition 3. To analyze the model correctness is equal to checking whether the model fits with the basic requirements described in definition 3. Building the reachability graph or tree of the model and using the following arithmetic can check the model correctness.

Step 0: all nodes in the reachability graph are marked as "non-checked". Traverse the reachability graph from $M_0.M \leftarrow M_0$
if the marker of M is "endpoint" then return false;
if component values of data places in $M > 1$ then return false, M marked as "checked"; insert M into queue.
while queue is not empty do take out the first element M from queue;

Step 1: for $M \neq M_{end}$ and all unchecked nodes M' linked with M

if marker of M' is "endpoint" return false
if component values of data places in $M > 1$ then return false
return false
M' marked as "checked"; insert M' into queue. return to step1;

Step 2: if the last component value of M_{end}=1 then

if all M marked as "checked" then return true;
else return false

3.2 Analytical Methods on Model Reliability

The reliability of service composition is influenced by the dependability of web service components. The analytical method used in the paper is based on the structural modeling. The reliability of web service composition is obtained by making use of the dependability of single web service components and

constructing Petri net models of web service composition. The following is about the analyzing method and its steps of service composition reliability.

1. The average executing time λ of web service composition is calculated according to the formula of equivalence performance of workflow-SPN model and the average executing time λi of single web service components.
2. Assume that λ takes on the exponent distribution, and the probability of the task being finished within $(0, t]$ demanded by customers can be worked out .

3.3 Security Analysis of Web Service Composition

The security analysis of web service composition mainly focuses on the ability of offering effective services when service composition faces all kinds of threats from networks. The security probability is adopted to measure the security of service composition. Firstly, build repairable model of invalid service; then, build the model's reachability tree and Markov chain; lastly, compute the steady probability, and the steady probability of initial marking is the security probability.

The repairable model of invalid services is used to describe the changing process of the state of system from normal to invalid, from invalid back to normal. Fig. 1 describes several repairable basic structures when invalid services occur; The basic structures of repairable model include sequence structure, concurrent structure and selective structure. As to the selective structure, because in one service composition flow only one branch can be selected at one time and service failures lead to system functional failures, the selective structure is equivalent with the sequence structure.

4 Credibility of Literature Retrieval and Payment System

In this part, the method of credibility analysis of SPN service composition is applied in a specific example of service composition. In the composition, mobile payment and electronic literature retrieval are combined. The process integrates
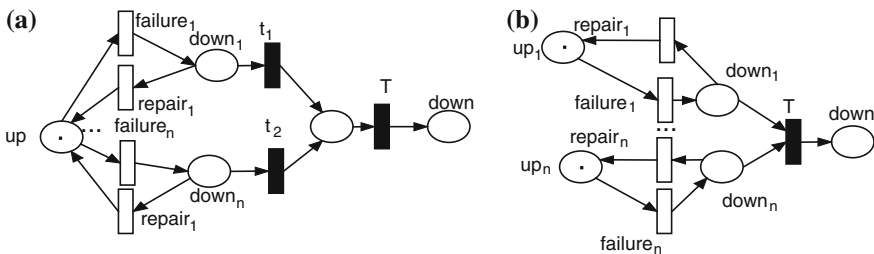

Fig. 1 Basic structures of repairable system. **a** is sequence structure **b** is concurrent structure
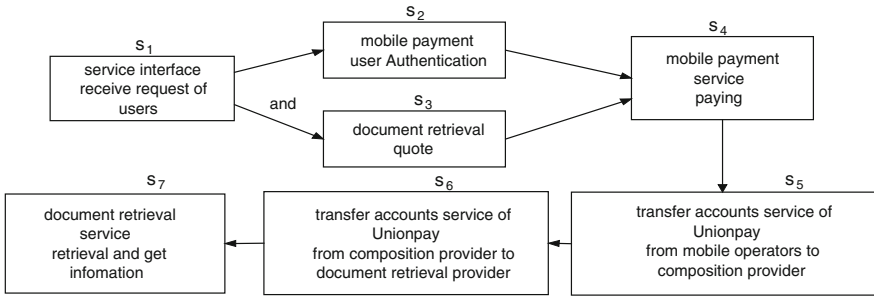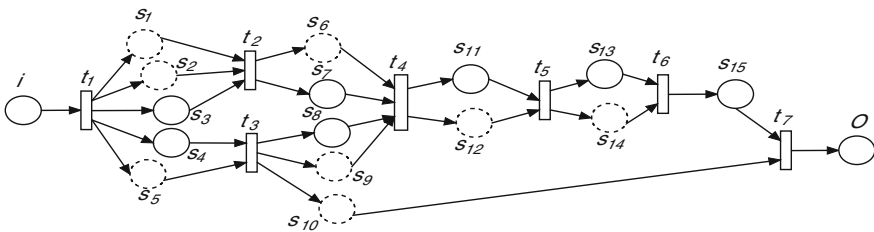
Fig. 2 Web services composition flow

Fig. 3 Web services composition model based on Petri net

resources provided by three business organizations. The literature research providers have electronic literature database, mobile operators have user-friendly payment services, and Union pay can provide transfer services for funds between enterprises. The flow of this web services composition is shown by (Fig. 2).

In Fig. 3, solid line cycles stand for control places; dotted cycles are symbols of data places. Transition means single web service of web service composition with the average delay time of service execution attached to it. The implement rate of transitions in the model is given in Table 1.

After the construction of the reachable tree, according to algorithm in 3.1, the model is correct, and can meet the user's demands. Calculate the execution time of web service composition. The average execution time of composition service is 148. The probability fulfilled in given time of service composition can be worked out according to the above-mentioned exponential distribution. Construct

Table 1 Implementing rate of transition in the model

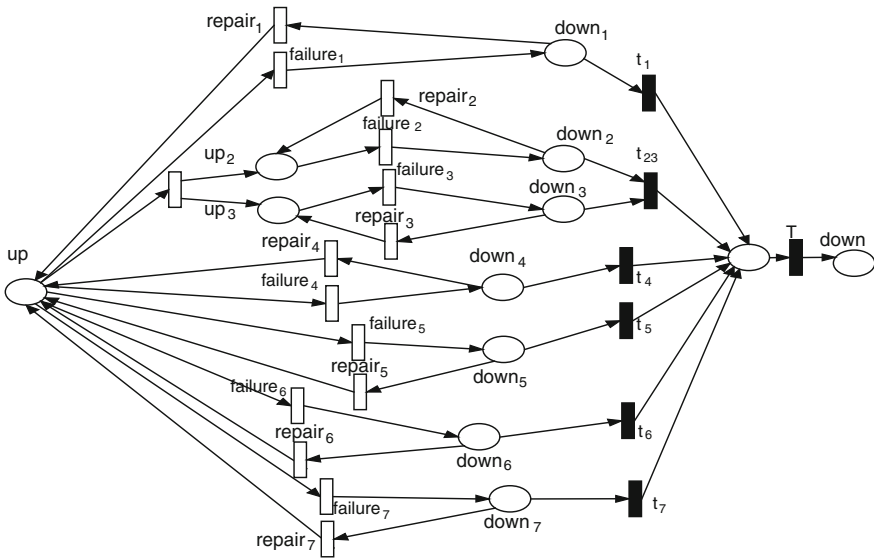| Transition	Implementing rate	Transition	Implementing rate
t1	1/8	t5	1/30
t2	1/30	t6	1/30
t3	1/10	t7	1/40
t4	1/15		

Fig. 4 Web services invalid model

Table 2 The invalid rate and repaired rate of transitions

Transition	Rate	Transition	Rate
Failure1	1/200	Repair1	1/200
Failure2	1/100	Repair2	1/100
Failure3	1/100	Repair3	1/100
Failure4	1/500	Repair4	1/100
Failure5	1/500	Repair5	1/100
Failure6	1/500	Repair6	1/100
Failure7	1/500	Repair7	1/100

repairable model of service failure of web service composition model, shown by Fig. 4; in Table 2, the invalid rate and repaired rate of transitions are given. The reachable graph, as is shown by Fig. 5, is established for the model simplified and steady probability is calculated. The Markov chain corresponding with the model is shown as Fig. 6.

The steady probability of the state M_1 is the probability of the system's ability to resume when facing vicious attacks, i.e., security probability. The security probability in this example is 0.4, which proves that the security performance of the service composition is not good enough.
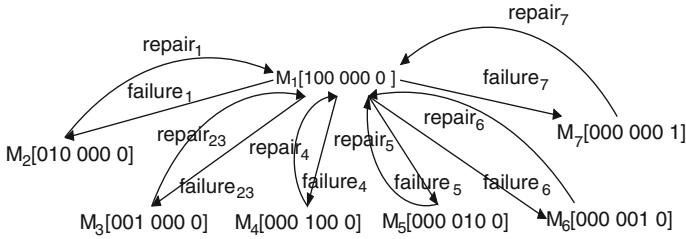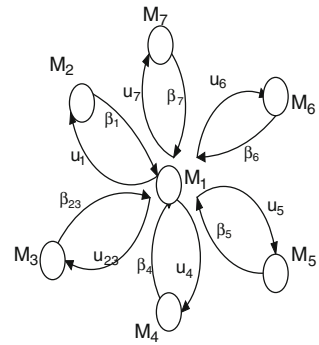
Fig. 5 Reachable tree of the model

Fig. 6 Markov chain

5 Summary

The indexes of correctness, dependability and security can be obtained by the method on web service composition credibility proposed in the paper so that a quick evaluation is available in terms of the providers' or users' credibility on web service composition. Some problems remain unsolved, such as the state explosion problem in SPN based modeling. It is hoped that problems as such can be removed in the future researches.

References

1. Cui H, Ying S (2010) Review of semantic web service composition. Comput Sci 37:21–25
2. Peiyun Z, Yaming S (2007) Research on dynamic web services composition. Comput Sci 34:4–7

3. Li X, Chen L (2009) Security study of web services composition. Appl Res Comput 26:1524–1527
4. Li J, Zhu X (2010) Fuzzy Petri net and its application in web services trust evaluation. Appl Res Comput 27:573–576
5. Zhao J, Liu L, Wu X (2008) A method of dynamic safety analysis for system based on Petri net. J Naval Univ Eng 6:9–12
6. Zhong D, Qi Z (2006) A Petri net based approach for reliability prediction of web services. Proc OTM'06, New Orleans
7. Yin YY, Li Y (2009) Determining on consistency and compatibility of web services behavior. Acta Electronica Sci 37:433–438
8. Zhao H, Sun J (2010) A methodological study of evaluating the dependability of SOA software system. Chin J Comput 33:2204–2209
9. Algirdas A et al (2004) Basic concepts and taxonomy of dependable and secure computing. IEEE Trans Dependable Secure Comput 1:11–13
10. Avizienis A et al (2001) Fundamental concepts of dependability. LAAS-CNRS, N01145, 2001, 4.
11. Lin C (2006) Research on network dependability analysis methods based on stochastic Petri net. Acta Electronica Sci 34:322–330
12. Lin C (2005) Stochastic Petri nets and system performance evaluation (2nd ed). Tsinghua University Press, Beijing

Based on DNA Self-Assembled Computing to Solve MH Knapsack Public Key Cryptosystems of the Knapsack Problem

Jing Liu and Zhixiang Yin

Abstract DNA self-assembly is a hierarchical build-up of complex assembly body; it is also a very important model in molecular computing. Cryptography problem not only has theoretical significance, but also has a very wide range of applications in national economy and other fields. We will use the way of self-assembly of DNA computing to solve the knapsack problem in the MH knapsack public key cryptosystem.

Keywords DNA self-assembly model · MH knapsack public key cryptosyste

1 Introduction

In 1994, Adleman successful use of DNA molecular biochemical reaction, the completion of the seven vertices have to Hamilton path problem solving [1]. DNA computing was born, it has opened a new computational science field, it has great theoretical and practical significance.

1.1 Research Status of Self Assembly of DNA Computing

DNA self-assembly model is a kind of important biological calculation model. It refers to molecules in thermodynamic equilibrium conditions, through chemical bond as combined to form stable structure spontaneously, complex and orderly, it

J. Liu · Z. Yin (✉)
Anhui University of Science and Technology, Huainan 232001, China
e-mail: zxyin66@163.com

J. Liu
e-mail: xingongliujing@126.com

has some special functions of molecular aggregates or the supramolecular structure of the process [2]. It represents a kind of the essence of DNA computing, known as the "bottom up" method. This method can be divided into two steps: at firstly, the way of synthetic chemistry is used to construct a single molecular; the secondly, the single molecules assemble into large structure. Without any external conditions for the intervention, DNA self-assembly algorithm can finally make small tiles forming precise large and complex structure. Seeman was firstly proposed by using DNA molecules self-assembly structure of Tile, he uses a DX Tile structure to establish many complex algorithm structures. The two dimensional self assembly models, is based on Tile theory by Wang et al., Winfree calls it "the Tile self-assembly model". In 2004, Rothemund et al. use DX Tile structure to realize a dimension cellular automata by the experiment, then they analysis and proof that the DNA Tile self-assembly structure realize the feasibility of any cellular automata, and self-assembly DNA calculations can realize the cellular automata XOR operation. In 2006, for the first time Rothemund use DNA origami technique, according to the sequence of DNA Tile were coded appropriate, he successfully constructed various nanoscale structure. In 2007, Brun proposed addition and multiplication model based on Tile self-assembly model. In addition, the Tile self-assembly model can effectively solve the combinatorial optimization problems, such as satisfiability problems, etc. In 2008, Brun was proposed for solving uncertain self-assembly algorithm of satisfiability problem.

The characteristic of molecular self-assembly is highly parallelism. A lot of different DNA Tile molecular bonded with the intermediate result, at the same time, they hold together. It is that the great advantage of DNA computing. But the different molecular would be self-assembly body parallel, then different self-assembly body seek test combination, and make the solution be better; Molecular self-assembly is a bottom-up complex assembly body. On the other hand, it allows intervention appearing in any steps, Due to the constraint conditions and influence, its growth was inhibited to other's direction. DNA, which has large scale parallel resistance, high capacity of the storage density and low energy consumption, is being widely development and research.

1.2 Research Status of DNA Code

The traditional password with the development of the electronic technology have great development in the 20th century, it is also widely used password system in currently; Quantum cryptography, firstly appeared in 1969, Stephen Wiesner puts forward "Conjugate coding" (Conjugate coding) concept. In 1984, Bennett. Chares H., Brassard. Gille based on this idea, proposed BB84 protocol. It based on quantum system coding scheme and key distribution protocol. Quantum cryptography was widely developed in recent years. But there are many problems to be solved in application; The DNA code is proposed after DNA computation, at present, it has become a new research hotspot of international code.

DNA password base on DNA as the carrier, with modern biological technology as a tool, using DNA molecular's characteristics and the advantage of DNA computing to realize encryption, decryption, certification, signature and other functions in Cryptography. At present, there are many relevant literatures about DNA code, such as Celland, he successfully put the famous sentence "June 6 invasion: Normandy" hidden in DNA micro sites. He realized to hide the information by using DNA as a carrier; Gehani developed the DNA code by providing people with further research in 2000; There are many relevant literatures; all of them have the different aspects of the DNA code [3–6].

2 The MH Knapsack Public-Key Cryptosystem

With the RSA public key, it is called one of the two famous public key systems. In the MH knapsack public-key cryptosystem we use 0–1 knapsack problem, namely: if we set b, a_1, a_2...a_n as some positive integers, make the equation has binary solution: $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b(x_i = 0 \text{ or } 1)$?

2.1 The Principle of the Backpack Password System

Set. $a_1, a_2...a_n$, it is a super increasing sequence, and $a_k > a_1 + \cdots a_{k-1}$, $k = 2, 3, \ldots, n$. Set. $b_1, b_2...b_n$, it is a common sequence, and $a_i, b_i \in Z_n, i = 1, 2, \ldots n$,

Set. $P = \{x_1a_1 + x_2a_2 + \cdots x_na_n | x_i = 0, 1, i = 1, 2, \cdots n\} = <a_1, \ldots a_n>$
$C = \{x_1b_1 + x_2b_2 + \cdots x_nb_n | x_i = 0, 1, i = 1, 2, \cdots n\} = <b_1, \ldots b_n>$,

Let T :P → C, T: $x_1a_1 + x_2a_2 + \cdots x_na_n \rightarrow x_1b_1 + x_2b_2 + \cdots x_nb_n$, so that $T(x_1a_1 + x_2a_2 + \cdots x_na_n) = x_1b_1 + x_2b_2 + \cdots x_nb_n$.

It is a reversible replacement, and its function is hiding the super increasing sequence, then they will become another sequence $b_1, b_2...b_n$. We will open them. For a given plaintext $M = m_1m_2...m_n$, $m_i = 0$ or 1. The encryption algorithm is $c = b_1m_1 + b_2m_2 + \cdots + b_nm_n$. For code analyzer, they will solute the general knapsack problem $b_1, b_2...b_n = c$.

2.2 Based on the Self-Assembly DNA Computing Solving Knapsack Problem

The principle of molecular self-assembly is that the program is spontaneous. The initial seeds will self-assemble into super-molecular [7]. DNA self-assembly calculation process has high parallelism. This kind of parallelism is divided into

global parallelism and local parallelism. Among them, the global parallelism is that every super structure contains information, the information represent different calculations; Local parallelism is refers to the individual super structure's growth in many local places at the same time.

The process of self-assembly DNA computing: It codes for the basic Tile sequences. This process divided into molecular recognition, molecular growth and growth termination. The process of molecular recognition is that the basic molecular selective match with other molecular; The process of growth process is that the basic molecular or intermediate state molecules complete the self-assembly process according to sequence or rank, this process fully embodies that the self-assembly have synergistic effect and nonlinear behavior; And the ending of the process is very important, if this process doesn't exist, self-assembly body will growth indefinitely, so we must set up certain constraints to set its termination in the practical process. The process is shown in Fig. 1.

Based on the DNA self-assembly model solving 0–1 knapsack problem, we build a model about the self-assembly of uncertain algorithm solving multidimensional bounded knapsack problem. There are several subsystem, including uncertain subsystem, multiplication subsystem, copy subsystem, addition subsystem and comparative subsystems, and then we use them to solve the problems and analyze the complexity of the algorithm. It shows that the algorithmic is feasibility and effectiveness. The theoretical analysis shows that we proposed algorithm using self-assembly technology, after constructing the initial chain, we needn't manual intervention, and self-assembly body began to growth and parallel execution operation. All the process has simple interpretation and low error.

The 0–1 knapsack problem can be described as: there is a backpack and n items. The backpack's capacity is C. The value of i item is v_i and the weight is w_i. Which items into backpack will make the total value of the largest if the weight of these objects does not exceed the capacity of the backpack? Namely:

$$Max\, z = \sum_{i=1}^{n} v_i w_i \, st. \sum_{i=1}^{n} w_i x_i \leq C \, x_i \in \{0, 1\}, i \in \{1, 2 \ldots n\}$$
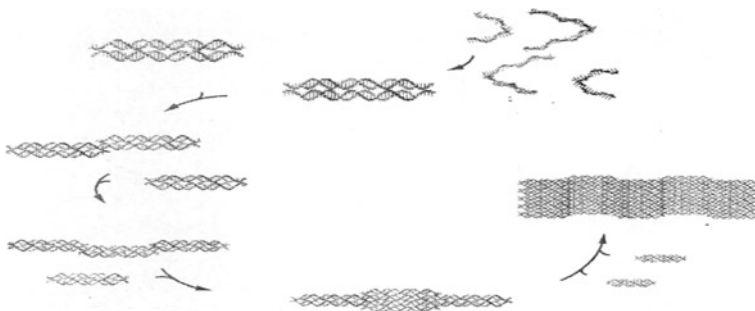

Fig. 1 DNA tile self-assembly process

Based on solving NP problem self-assembly model by Brun, now it is improved: firstly we construct a system for identifier and selection (construct seed configuration, uncertain subsystem and a addition subsystem); secondly we need calculate the value of all goods; thirdly, we use constraints to exclude those infeasible method, it is required to verify the value of each combination and find the relationship between the them and C (compare subsystem); Finally, we use self-assembly process foe more time to find the optimal solution. In a word, this process contains these steps and four systems, uncertain subsystem, add subsystem, copy subsystem, comparative subsystem.

Following a brief introduction about the system:

1. Add subsystem

This subsystem make different variables and their corresponding multiplication system do add operation. In the top of Fig. 2 used to calculate the weight of all the goods into the backpack; the middle one is used to calculate the corresponding value of the selected items. The two system of Tile are same, so that v_i and w_i are corresponding. The top tile input from the left and the below, the output from above, on the right is carry value end. For example, One addend is $(011001)_2$, another one is $(1111)_2$, the result is $(101000)_2$, As shown in the bottom of figure.

2. Copy subsystem

The assembly process, copy subsystem is not used to participate in any operation, it is just responsible for transmitting from low-end to top-end, in this process, and the data does not change. For example, the low-end data is $(011011)_2$, this data is transmitted to top-end, and $(1011)_2$ of its right does not any change, its task is that the low-end data is transferred to the top, the right data does not participate in the operation. As shown in Fig. 3. The first one of Fig. 3 is composed of four Tiles, which is called copy subsystem, their value in the top and bottom are the same, and the data in the bottom will be transmitted to the top, the left and the right only has a function of connection. They contain all the combination of 0 and 1. The next figure says a simple example. That is to say, when the item is not selected, its value is 0 or 1.

3. Uncertain subsystem

Using self-assembly method to solve problems, add subsystem and copy subsystem must be used, but which one should we choose is not sure. Namely, we choose among them as a certain system are random in the self-assembly process, so the value of x_i is random. If this kind of random is the true random, the set $\{x_i\}$ will probably take all the value, and get all of the possible solutions. It requires a recognition process, for example, if the end of uncertain Tile connection is "?" It means that is uncertain subsystem; when its left side shows " $\sim$ ", its right side will connect addition subsystem; if shows "x", it means that the right side will connect copy subsystem.
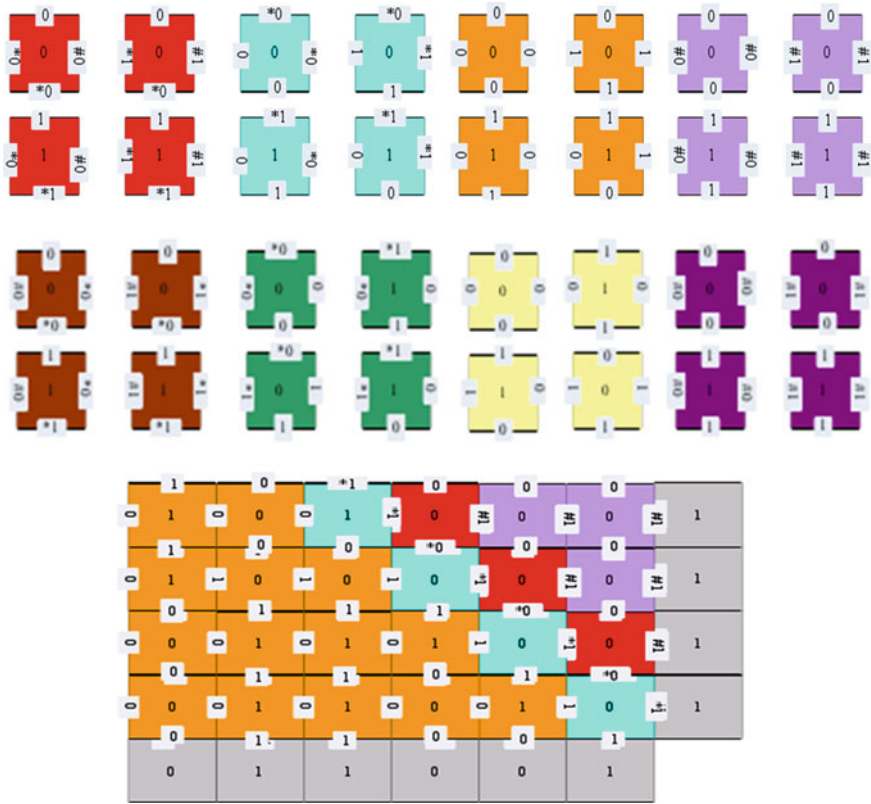
Fig. 2 Addition subsystem

4. Comparative subsystem

The system is mainly used to determine the size of the inputted positive. At first, according to the two numerical binary forms, the data in the low-end is compared; the dates in the corresponding position will be compared and marked. Then the relationship between two numbers was determined. The 0–1 knapsack problems as an example, it needs to compare with all the weight of the goods in the backpack. Here, all the numbers are binary digits of positive integer and they all have the same digits. At the end of the comparison, if it appears "$\sqrt{}$", it means the weight of the selected goods less than the capacity C of the knapsack, and this kind of method is feasible. But it may be not the best method and it still needs the further calculation; if it appears "x", it means that is not feasible. As shown in Fig. 4. We can see an example: $a = (101001)_2, b = (100110)_2$, Tile tells us ">",the result of the Comparison is a > b.

Based on the self-assembly DNA computing huge parallelism. We can test the results at the same time. The DNA self-assembly model, need to perform the following steps in solving knapsack problem: (a) the uncertain subsystem generate
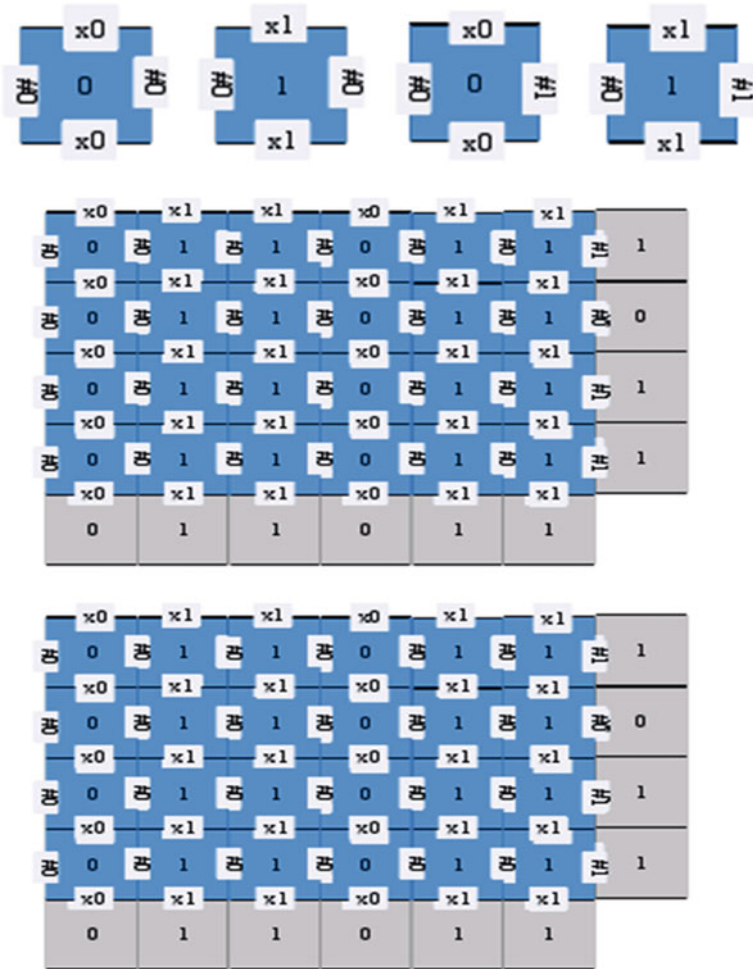
Fig. 3 The four tile of the replication subsystem and an example

the vector $\vec{x}$, it contains all the possible combinations of 0 and 1, then all the possible solutions of the objective function are found. (b) All the possible solutions are checked by using the constraint condition in formula at the same time, unsatisfied solution will be eliminated, and the feasible solution will be kept. The corresponding coefficient's produced and the variable will be calculated in multiplication subsystem, the addition subsystem can executive additive operation of the multiplication operation's results, comparative subsystem can ensure that whether the result is less than or equal to the upper bound of the backpack's capacity, and at the same time, copy subsystem will transmit information to the upper Tile. And it repeats the process until you get rid of all the solutions; (c) All the feasible solution will be compared, we will acquire the optimal solution or
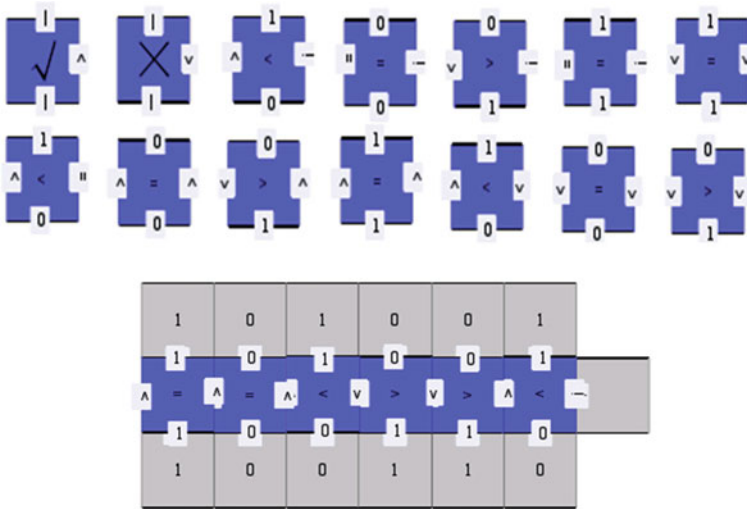
Fig. 4 Compared tiles and an example

approximate solution. This process requires a course of biological operation, gel electrophoresis operation can separate out the different length, it will be the output results chain, then the report chain contains special marked as primer and extract the chain, finally the technology of the PCR amplification amplification and DNA sequence of standard determination are used, and the result will be read.

3 Conclusion

R. C. Merkl and M. E. Hellman use the puzzle of solving knapsack problem to construct the difficulty of the backpack password system, DNA self-assembly computing can solve this puzzle efficiently. Its time complexity and space complexity are all $\theta(n)$, The species of Tile are $\theta(1)$ [8]. Of course, the self-assembly DNA computing can also be used in a lot of encrypting and deciphering algorithm of Cryptography, with the effort of the previous research and the future DNA computing researchers, the DNA self-assembly computing will be used in more fields and solve more practical problems.

References

1. Adleman LM (1994) Molecular computation of solution to combinational problem. Science 266:1021–1024
2. Yan MF, Zhao Z, Cui L (2010) DNA folded technique research progress. Appl Chem 27:126–132
3. Gehani A, LaBean T, Reif J (2004) DNA based cryptography. Aspects Mol Comput, p 34–50
4. Xiao G et al (2006) New field of cryptography: DNA cryptography. Chin Sci Bull 51:1413–1420
5. Lu M et al (2007) Symmetric-key cryptosystem with DNA technology. Sci Chin Ser F: Info Sci 50:324–333
6. Lai X et al (2010) Asymmetric encryption and signature method with DNA technology. Sci Chin Info Sci 53:506–514
7. Seeman NC (2003) Biochemistry and structural DNA nanotechnology : an evolving symbiotic relationship. Biochemistry 42:7259–7269
8. Adleman L, Cheng Q, Goel A et al (2002) Combinatorial optimization problems in self-assembly. Thirty-Fourth annual ACM symposium on theory of computing, Montreal, Quebec, Canada, pp 19–21

Image Representation and Recognition Based on Directed Complex Network Model

Ying Chen, Jin Tang and Bin Luo

Abstract Image structure representation is a vital technique in the image recognition. A novel image representation and recognition method based on directed complex network is proposed in this paper. Firstly, the key points are extracted from an image as the nodes to construct an initial complete undirected complex network. Then, the k-nearest neighbor evolution method is designed to form a series of directed networks. At last, the feature descriptor of the image is constructed by concatenating the structure features of each directed network to finally achieve image recognition. Experimental results demonstrate that the proposed method outperforms the traditional methods in image recognition and can describe the structure of images more effectively.

Keywords Image recognition · Directed complex network · Dynamic evolution · k-nearest neighbor · Feature extraction

1 Introduction

Image representation plays a vital role in computer vision and image processing. There are various categories of image representation. Texture features-based methods [1, 2], edge and contour-based methods [3, 4], and the key points-based

Y. Chen · J. Tang (✉) · B. Luo
School of Computer Science and Technology, Anhui University, Hefei 230601, China
e-mail: ahhftang@gmail.com

Y. Chen
e-mail: chenying0557@126.com

B. Luo
e-mail: luobin@ahu.edu.cn

J. Tang · B. Luo
Key Lab of Industrial Image Processing and Analysis of Anhui Province, Hefei 230039, China

methods [5–7] are some of them. The key points-based methods are one of the most popular categories owing to its lower computational complexity. However, it is still a challenge to the traditional structural description methods due to the ubiquitous image noise which will result in the extraction error of key points [5, 7].

Recently, complex network theory is becoming more and more attractive [8, 9] and has been widely applied in pattern recognition. Backes et al. [1] analyzed the image texture features and established a complex network model to achieve the identification of texture image; Backes et al. [3, 4] established a complex network model on the shape contour edge points to achieve the shape feature description. In addition, Tang et al. [10] achieved the feature representation of structure graphs using complex network model. However, the above models were all established based on the undirected graph. In fact, the directed graph often contains more abundant structural information. Therefore this paper gives a novel directed complex network representation model. Furthermore, the image feature extraction will be conducted through the complex network features under the k-nearest neighbor evolution. Benefited from the statistical features, this image feature description has the advantages of favorable stability and strong anti-noise ability. Compared with the undirected complex network model, the directed complex network model contains more abundant information and can describe the image structure more effectively, therefore can produce better recognition results.

2 Preliminary Work

In this section, the basic concepts of complex networks that are related to our work are briefly recalled.

2.1 Representation of Complex Network

Complex network can be represented as graph $G = <V, E>$, where node $v_i \in V$ and edge $(v_i, v_j) \in E$ denotes the pairs of neighboring vertices. Each edge $(v_i, v_j) \in E$ has a associated weight $w((v_i, v_j))$, which is a non-negative measurement of the dissimilarity between neighboring nodes v_i and v_j. It is noted as undirected network if node pairs (u, v) and (v, u) represent the same edge, otherwise known as the directed network. It is noted as a weighted network if each edge of the network has been allocated a weight, otherwise known as the unweighted network. We focus on the directed weighted network in this paper.

2.2 Measurements and Dynamic Evolution

Each complex network has some specific measurements to characterize and describe the topological features of the complex network. Therefore, the comparison and analysis of complex networks can be achieved by computing and analyzing these measurements [3]. A complete survey of measurements can be found in Ref. [9]. The six measurements for complex network description used in this paper are the maximum degree, average degree, entropy, energy, average joint degree and average shortest path length [3, 9].

Dynamic evolution is another important characteristic for complex networks. The trajectories of the measurements for the complex network can be obtained by concatenating the measurements of networks extracted from different stages of the evolution process [3, 9] which provides a more comprehensive way to analyze and discriminate networks. Different network structures can be identified by comparing these various measurement trajectories. If their measurement trajectories evolve in a similar way, these networks may have similar structures [3, 9].
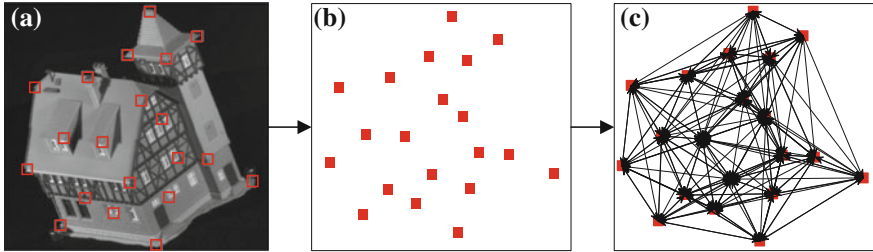
3 Image Representation Using Directed Complex Network

Based on the above features of complex networks, the key points are modeled into the directed weighted network using complex network. Then the network topology features are extracted under the dynamic evolution to achieve the image feature extraction.

3.1 Image Representation

As inspired by the work [3], using complex network model, a representation of N key points should be generated as graph $G = <V, E>$. Each key point of the image is represented as a vertex in the network, which is constructed by a set of non-directed edges E binding each pair of vertices. The weight w_{ij} for the edge connecting vertex i and j is calculated by Euclidean distance. Therefore, the network can be represented by an $N \times N$ weight matrix W, normalized into interval [0, 1].

As discussed in the work [3], the set of edges E connects all the vertex pairs in the network initially. This network is a complete undirected graph without any topological property. Therefore, it is necessary to transform the initial network into a series of complex networks [3]. Fig. 1 shows the representation process of an image using the directed complex network model.

Fig. 1 Image representation with the directed complex network. **a** Initial image. **b** Key points extraction. **c** Directed complex network representation

3.2 Dynamic Evolution

Dynamic evolution is a vital characteristic for complex networks. Backes et al. [3, 4] exploited the edge weight threshold (EWT) evolution method to carry out this transformation, but the sub-networks under this evolution are undirected graphs. It is noted that directed graphs always contain more abundant information. Therefore a novel k-nearest neighbor (kNN) evolution is proposed in this paper to produce the directed sub-networks. Given the initial network G_0, the evolution process is defined as follows:

$$G_k = \delta(G_0, k) = \begin{cases} w_{ij} = 1 & \text{if } j \in kNN(i) \\ w_{ij} = 0 & \text{otherwise} \end{cases} \tag{1}$$

where $k = 1, \ldots |V|\text{-}1$ and $kNN(i)$ denotes the k nearest neighbors of the node i. Fig. 2 shows the image and the dynamic evolution process of directed complex network, where the left two subfigures denote the initial image and the corresponding key points from top to down and the right six subfigures denote the directed sub-networks with $k = 3, 5, 7, 9, 11, 13$ respectively from left to right and top to down.

3.3 Feature Extraction

Based on the kNN evolution method, the six features used in this paper are extracted. Since the out-degree of all the networks is identical (equal to k), only the in-degree is taken into account. For each network, the feature vector ϕ can be expressed as follows:

$$\begin{aligned} \phi &= [F_1, F_2 \ldots, F_{|V|-1}] \\ F_i &= [d_\kappa(i), d_\mu(i), H(i), E(i), P(i), l(i)] \end{aligned} \tag{2}$$
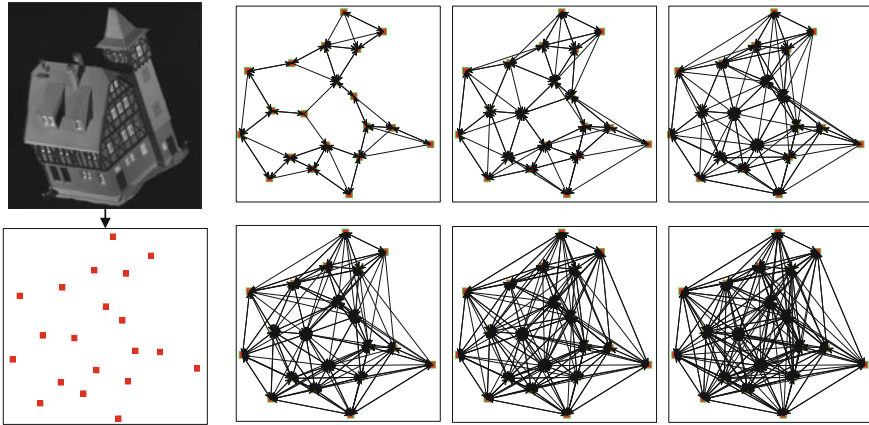
Fig. 2 Image and the dynamic evolution process of directed complex network

where $d_\kappa(i)$, $d_\mu(i)$, $H(i)$, $E(i)$, $P(i)$, $l(i)$, $(i = 1, 2, \ldots |V| - 1)$ denote the features of the ith sub-directed network including maximum degree, average degree, entropy, energy, average joint degree and average shortest path length [3, 9, 10].

4 Experiments

To evaluate the performance of our method, experiments were conducted on two real-world images datasets, i.e., building sequence (CMU, MOVI and YORK) datasets and Columbia Object Image Library-100 (COIL-100) dataset [5, 10, 11].

4.1 Pattern Space Embedding

The object clustering experiment was conducted on the building sequence datasets. Thirty images were selected from different datasets in this experiment. The first class contained 10 images from CMU, which were obtained by rotating objects from the top. The second class consisted of 10 images selected from the MOVI database with the images being obtained by rotating in the 3D scene. The third class consisted of 10 images from the YORK that were acquired by rotating objects around the vertical axis in the 3D scene [7, 11]. The number of key points extracted from CMU, MOVI and YORK were 32, 80 and 41, respectively. k varies from 1 to 31 (in step of 1).

Firstly, the distance matrices based on the proposed kNN method comparing with the EWT evolution [3, 4] and Graph Edit Distance (GED) method [7] were computed to demonstrate the similarity between the pairwise images, as shown in
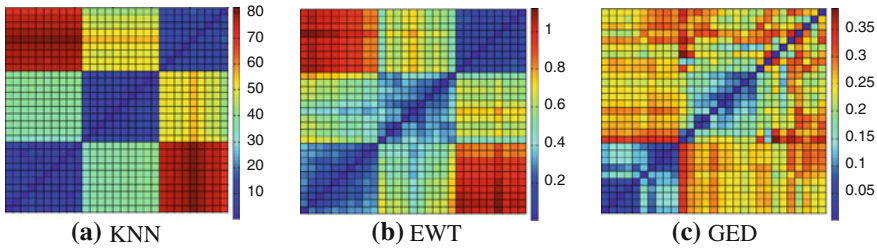
Fig. 3 Distance matrices with three methods on three sequence datasets

Fig. 3. The darker the patch is, the smaller the value in the matrix is. The comparison result shows that our directed kNN method (Fig. 3a) can distinguish the classes more clearly than the other two undirected methods (Fig. 3b, c). Furthermore, in our method, every image was significantly closer to its own class than to the other classes.

In addition, to compare the capability of capturing the potential dissimilarity among different images, Principal Component Analysis (PCA) and Multidimensional Scale Analysis (MDS) methods [7, 10, 11] are employed to implement pairwise data clustering with feature matrices and distance matrices respectively. The comparison results on different methods are shown as Fig. 4.

It can be seen in Fig. 4 that images of each class embedded with kNN (Fig. 4a, d) are more tight and distinct classes can be separated better in the embedding space compared with those in EWT (Fig. 4b, e) and GED (Fig. 4c, f)
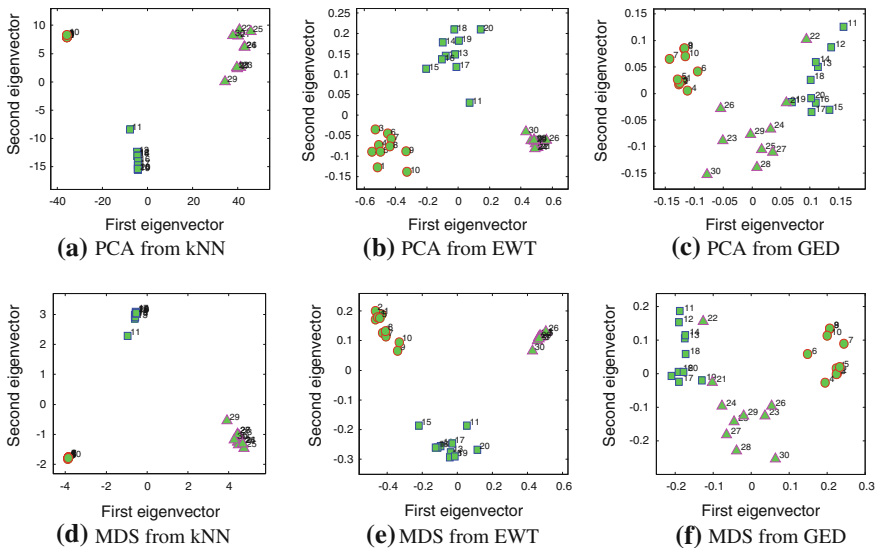

Fig. 4 PCA and MDS for each of distance matrices on three sequence datasets

methods [3, 7]. This demonstrates that our directed method contains more abundant structural information and thus can capture the potential dissimilarity among different images more effectively.
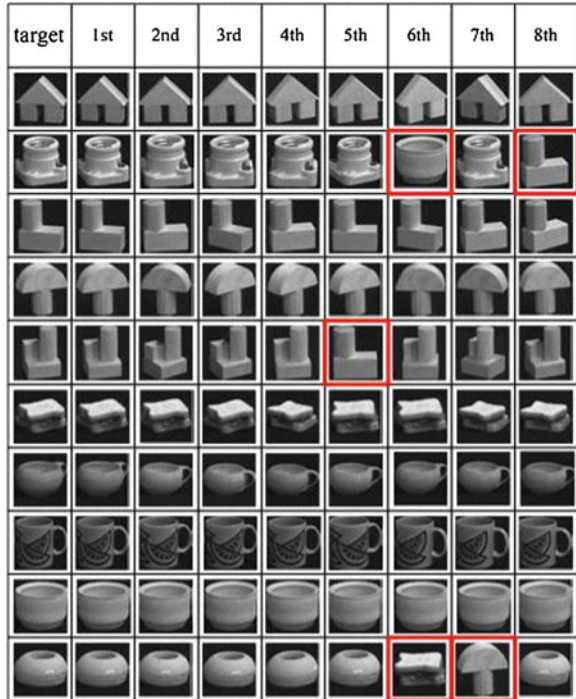
4.2 Image Retrieval

COIL-100 dataset [7] contains 2D viewing images for 3D objects of 100 classes. Images in each class were obtained by fixing a camera and rotating the object sequentially by 5° around a vertical axis. The graph structure of images in the same class may be completely different when the rotation angle is too large. Therefore, 150 images of 10 classes were selected from the COIL-100 to conduct retrieval experiment. In this section, the complex network model based on the kNN evolution method was employed, the GED and EWT methods [3, 7, 10] were selected for comparison. The number of Harris corners extracted from each image is 45.

Recall R and precision P are important indexes to measure the retrieval performance. R is defined as the ratio of the number of retrieved relevant images r to the number m of relevant images in the whole database, i.e. $R = r/m$. P is defined as the ratio of the number of retrieved relevant images r to the total number of retrieved images n, i.e., $P = r/n$ [12]. The precision-recall curve (PVR) can intuitively react to the performance of a retrieval system. The greater the area surrounded by the precision-recall curve and the two coordinate axes is, the better the performance of the system is [12, 13]. Figure 5 shows the comparison results, from which we can see that the kNN (Fig. 5 red curve) is superior to the GED (Fig. 5 blue curve) and EWT (Fig. 5 black curve) methods. This demonstrates that the kNN evolution is a more effective evolution method. Therefore, the

Fig. 5 Average retrieval precision-recall curve with different methods in COLL-100 database
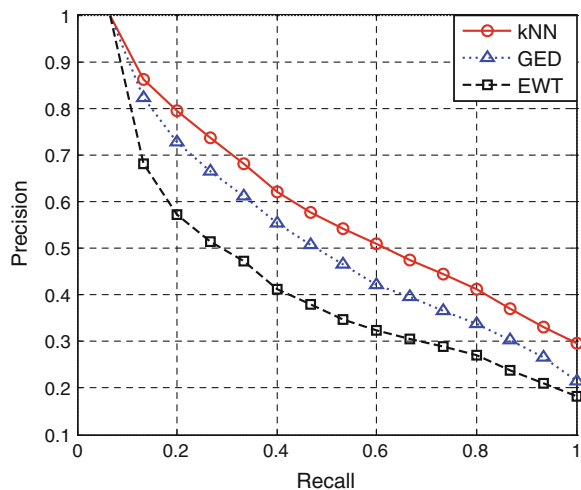
Fig. 6 Part of the retrieval results on COIL-100 dataset (Retrieval errors are marked by red rectangles)

performance of directed image representation based on kNN evolution method is competitive.

Fig. 6 shows some retrieval results. Here the first column is the target images and the remaining eight columns are the retrieval results.

5 Conclusion

A novel image representation and feature extraction method based on directed complex network model is proposed in this paper. The satisfying image clustering and retrieval experimental results demonstrate that the proposed method can describe image structure information more effectively compared with other undirected methods, and it also can better describe the structure features of complex networks. In our future work, we would like to integrate the content features of the images as well as statistical features of the networks to further implement image feature extraction.

References

1. Backes AR, Martinez AS, Bruno OM (2011) Texture analysis using graphs generated by deterministic partially self-avoiding walks. Pattern Recogn 44(8):1684–1689
2. Jianguo Z, Tieniu T (2002) Brief review of invariant texture analysis methods. Pattern Recogn 35(3):735–747
3. Backes AR, Casanova D, Bruno OM (2009) A complex network-based approach for boundary shape analysis. Pattern Recogn 42(1):54–67
4. Backes AR, Bruno OM (2010) Shape classification using complex network and Multi-scale Fractal Dimension. Pattern Recogn Lett 31(1):44–51
5. Bin L, Edwin RH (2001) Structural graph matching using the EM algorithm and singular value decomposition. IEEE Trans Pattern Anal Mach Intell 23(10):1120–1136
6. Jin T, Chunyan Z, Bin L (2006) A new approach to graph seriation. In: Proceedings of International Conference on Innovative Computing, Information and Control (ICICIC'06) (2006)
7. Xinbo G, Bing X, Dacheng T et al (2008) Image categorization: Graph edit distance + edge direction histogram. Pattern Recogn 41(10):3179–3191
8. Amaral LAN, Ottino JM (2004) Complex networks. Eur phy J B 38:147–162
9. da Costa L, Rodrigues FA, Travieso G et al (2008) Characterization of complex networks: a survey of measurements. Adv Phy 56(1):167–242
10. Jin T, Bo J, Chin-Chen C et al (2012) Graph structure analysis based on complex network. Digital Signal Process 22:713–725
11. Bin L, Wilson RC, Hancock ER (2003) Spectral embedding of graphs. Pattern Recogn 36(10):2213–2223
12. Xin S, Xiaojun W (2011) A novel contour descriptor for 2D shape matching and its application to image retrieval. Image Vis Comput 29(4):286–294
13. Xiang B, Bo W, Cong Y et al (2012) Co-transduction for shape retrieval. IEEE Trans Image Process 21(5):2747–2757

Finite Element Analysis and Computation of the Reinforcement Effect and Mechanism of the Ventilation Roadway U-Type Arch

Ning Zhang, Mingzhong Gao and Laiwang Jing

Abstract Though the U-type arch support is used widely in the environment of high earth pressure, damage rate still can not be controlled in some complicated condition. On the basis of U-type arch, the reinforcement structure established can improve the support intensity and control the deformation. The essay makes an adequate finite element analysis and computation on the U-type reinforcement structure by using numerical simulation software to test the validity of the reinforcement structure.

Keywords Reinforcement structure · Finite element analysis computation

1 Introduction

Throughout the transformation process of the coal mine roadway in our country, it can be seen that the support strength has been constantly reinforced and the support resistance also increased. A lot of research has been done at home and abroad on the controlling mechanism and techniques of high stressed roadway surrounding rock. The techniques developed include the high-resistance and extensible U-type steel wall backfilling technique, the new bolt grouting support

N. Zhang (✉) · L. Jing
College of Science, Anhui University of Science and Technology,
Huainan 232001, China
e-mail: zh_nin@sina.com

L. Jing
e-mail: lwjing229@163.com

M. Gao
College of Energy and Security, Anhui University of Science and Technology,
Huainan 232001, China
e-mail: mzgao@aust.edu.cn

technique [1, 2], and high-strength whole set bolt support technique and so on, which have been popularized and applied widely in coal mines [3–5], and have achieved remarkable economic benefit. It is worth mentioning that the U-type arch support has used widely in the roadway of the coal mine in Huainan-Huaibei area, to solve the problem of roadway deformation caused by soft rock and high earth pressure. However, single U-type arch support does not work well in the conditions where high-strength support is needed, so the essay makes an attempt to improve the support theory of U-type arch support through numerical analysis on the performance of reinforced U-type arch support by using ABAQUS software.

2 Project Background

Yuanzhuang coal mine has a history of fifty years, and its IV2 special ventilation roadway was bolting and shotcreting support originally, which was later changed to U-type arch support. Because the high pressure and the floor heave of the roadway, the U-type arch support has deformed. On the schedule of the coalmine, floor lowering and arch changing repair has been done, lowering floor nearly 180 m, changing arch about 40 m, with an engineering quantity of 220 m in total. Three-stage 29UU type arch support has been used in the repair work. The clear width of the roadway is 2.8 m, and the clear height 2.8 m. Steel fabric whole sealing to the top, the specifications of the steel fabric net is length × width = 800 × 400 mm, using 10 mm diameter rebar to weld the main reinforcement, and 6 mm rebar for the sub-reinforcement.

 The whole metal support can be seen as a symmetrical structure in accordance with the vertical symmetry axis of the roadway. In terms of the analysis approach of the symmetrical structure, it is allowed to take a half as the subject.

 As it is in bilateral symmetry, neither can the symmetrical section A moves on the left and right direction, nor can it twirl, so the left half is chosen as the subject of the research, which can be treated as a fix-end section, as shown in Fig. 1. According to symmetry, there is not a vertical constraint reaction in the constraint reactions of the fix end.

3 The General Situation of Finite Element Numerical Calculation Model

This simulation will adopt the contrastive approach to analyze the bearing capacity of the Yuanzhuang IV2 special return air course U-type arch support before and after the reinforcement.

 Taking full advantage of the symmetry, the whole model only makes a simulation of the right half of the prototype lengthways symmetrical plane, which
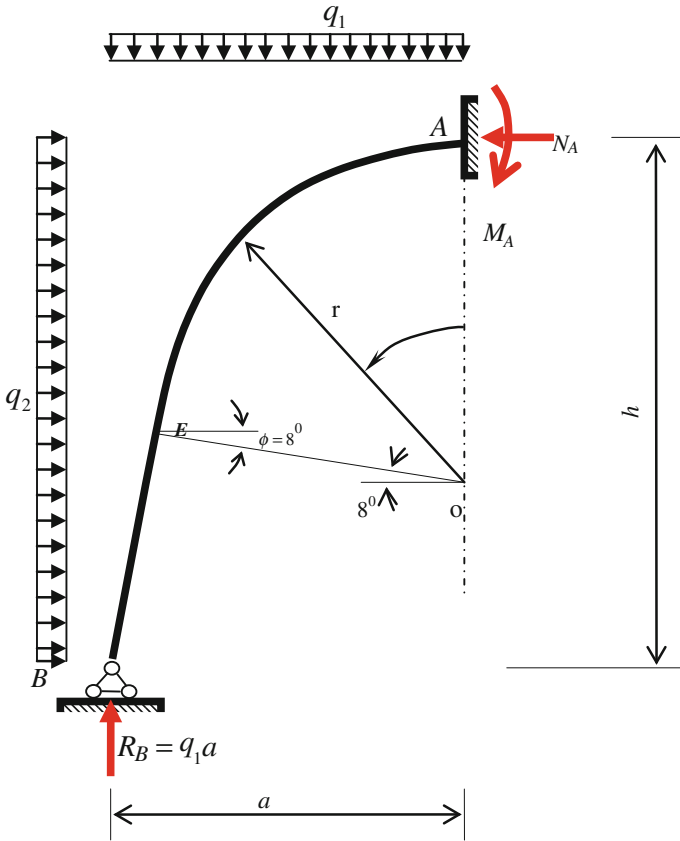
Fig. 1 Mechanics simplified model

makes full use of the computer resource. The reinforcement structure is added to one of the three-dimensional finite element models, and not to the other. The model without reinforcement structure adopts hexahedron element dissociation. Because the fracture surface of the U-type steel is complicated, which could not be meshed effectively by using the "split method" of the ABAQUS, so sweep method is adopted to mesh the whole model. The meshing model is shown in the diagram. There are 48,500 nodal points and 33,282 units, without warning and fault units. The analysis uses linear hexahedral elements of type C3D8R [7], which can ensure the accuracy of the analysis result. The reinforcement structure and the U-type steel support contact much. Face-to-face contact is taken as the analysis type. Because there is a small geometric displacement between the own reinforcement structure and the U-type steel support, the small slippage formula is used to calculate it, and the normal contact attribute adopted is set as hard contact. And because the hexahedron element cannot be used in the partition due to complexity
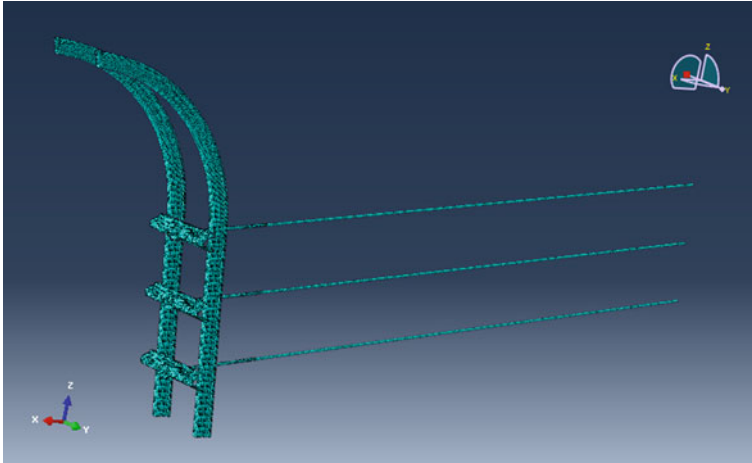
Fig. 2 Finite element model meshing schematic diagram

of the geometry of the reinforcement structure, tetrahedron element is adopted and to get more accurate pressure numerical value on the contact spot, C3D10 M element is used (Fig. 2).

4 The Summary and Result Analysis of the Reinforcement Structure

The U-type steel arch support has a strong capacity to withstand vertical earth pressure, but because the existence of the floor heave and its capacity to stand the side pressure, the damage to the arch support mainly results from the low intensity of the column leg to resist curving. Under high side pressure, the arch leg move inwards, causing changes to the whole pressure structure, resulting in a considerable fall of the supporting capacity and a final lose of support capacity. Therefore, the key to improve the support capacity is to restrain the deformation of the column legs, while the floor heave of the Yuanzhuang IV2 special ventilation roadway is in a serious condition, and though the column legs insert the floor to 20 cm, the inward movement can not be constrained effectively, so we used reinforcement structure to study its support strength.

4.1 Model Overview

Geometrical size

As shown in Fig. 3, the interval among the U-type arch supports is 600 mm, Using 29U reinforcement structure for mine 29U type steel for 750 mm, the supporting plane is $15 \times 150 \times 150$ mm square steel plane, the anchor cable is 6.3 m anchor cable for coal mine use. All the materials used are treated according to linear elasticity, and temperature expansion coefficient is added to the parameters of the anchor cable material.

The key of the model is the effect of the simulative reinforcement structure, using the cold contract method in ABAQUS to apply a pre-stress on the anchor cable element in the way of lowering the temperature to add the pre-pressure.

$\xi = \Delta T \times \alpha$, ξ—Strain; ΔT—Temperature difference; α—Coefficient of thermal expansion [8].

Boundary condition

The same as the structure without reinforcement, the boundary condition of the model is the displacement boundary condition: because the structure taken is symmetrical structure with a symmetrical boundary added to the model to constrain its displacement in the X direction; while according to the simplified mechanics model, add a displacement constraint in the Z direction to the bottom.

But the difference from the no reinforcement structure is the adding of the anchor cable and reinforcement structure [7]. The reinforcement structure and the U-type steel support contact much. Face-to-face contact is taken as the analysis type. Because there is a small geometric displacement between the own reinforcement structure and the U-type steel support, the small slippage formula is used to calculate it, and the normal contact attribute adopted is set as hard contact.
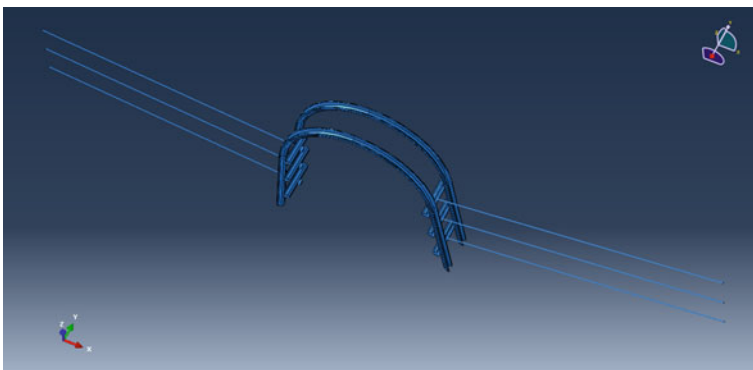
Fig. 3 The schematic diagram of the geometric size of the reinforcement structure adding model

The constraints on the anchor cable are fix-end constraint on the head, and Z direction constraint on the anchorage segment (Fig. 4).

Material attribute

All the U-type steel uses the Q275 parameter as well, and in the simulation process, all the materials are treated according to linear elasticity. And the anchor cable is treated as well, adding the Expansion of $1E - 5$.

Loading procedure

The reinforcement process of the model involves two procedures: (1) lower the temperature of the anchor cable to 50 and add a 10 Mpa pre-stress; (2) using the amplitude technique to add surrounding rock load to the U-type steel support gradually to the required magnitude (0.15 Mpa).

4.2 The Analysis of the Simulation Results

Mises stress result

In the first procedure of lowering temperature to 50, it can be seen that the anchor cable pre-stress is 10 MP or so, the result meets the requirements. The correctness of adding procedure (1) is proved (Fig. 5).
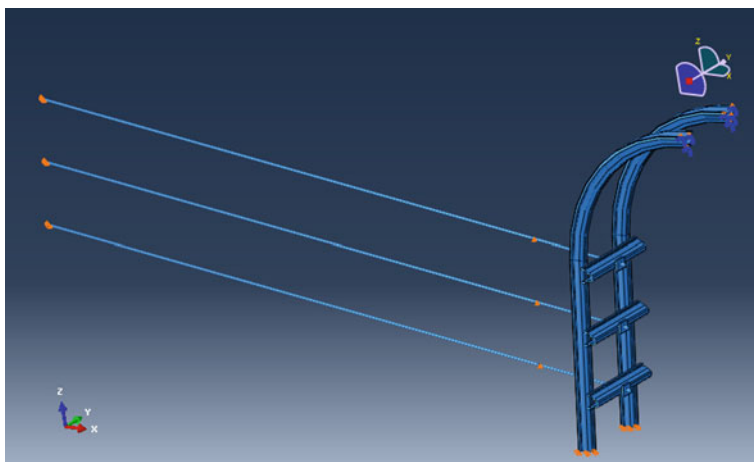
Fig. 4 The boundary condition of the model
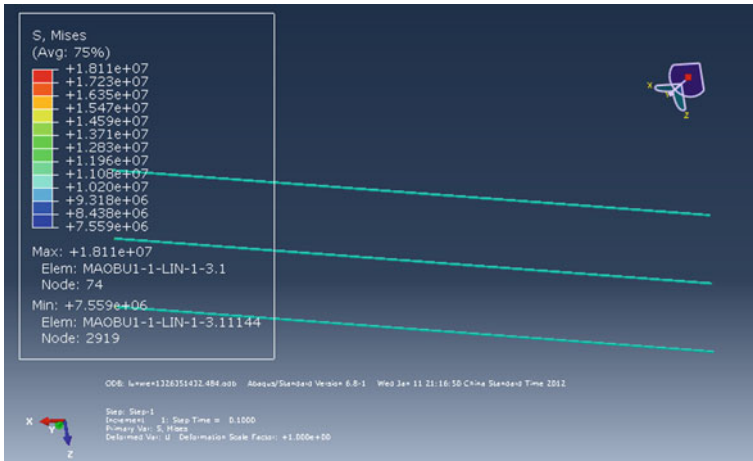
Fig. 5 Anchor cable pre-stress diagram
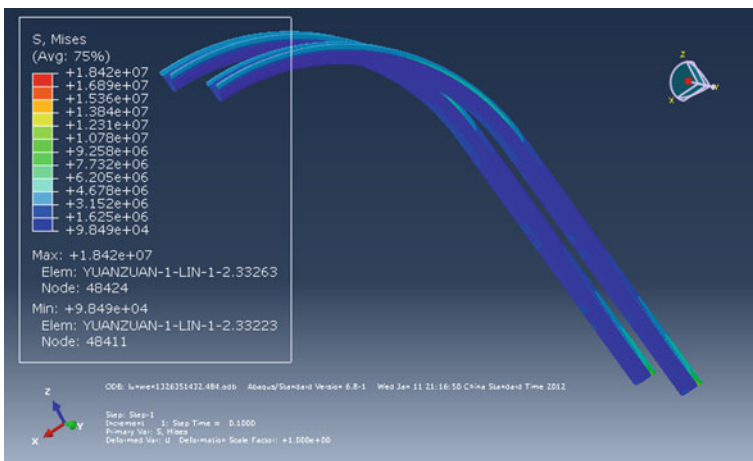
Fig. 6 U-type steel support mises stress cloud diagram

U-type steel support Mises stress cloud diagram, as shown in the diagram, the largest stress is 18.42 Mpa, the spot is the stress focus area on the bottom, while the whole distribution change is mild (Fig. 6).

5 Conclusion

It shows that the reinforcement structure lowers the whole structure stress. Before reinforcement the highest stress tested is 525.3 Mpa (in the range of limit rupture of the tensile resistance of Q275 steel), and the lowest stress is 6.625 Mpa. While after the add of the reinforcement structure, the highest stress is only 183 Mpa (not achieve the yield strength of Q275 steel), and the lowest stress is 1.013 Mpa. There is a very clear strain decrease, which has achieved about 5 times. It can be seen that the reinforcement structure can strengthen the U-type arch support, which is worth popularizing in a wide way among the coal mine under high earth pressure.

References

1. Chun'an Y (2000) Internal force calaulation of U-supports considering yielding. Chin J Geotech Eng 22(5):604–607
2. Xie W (2010) Structural stability of U-steel support and its control technology. Chin J Rock Mech Eng 29(9):604–607
3. Shi Y, Zhou Y (2006) ABAQUS finite element analysis of detailed examples. Machinery Industry Press, Beijing
4. Chun'an Y (2002) Stability analysis of U-steel yieldable support. Chin J Rock Mech Eng 21(11):1672–1675
5. Zhang N, Gao M (2004) High-strength and pretention bolting support of coal roadway and its application. J China Univ Min Technol 33(5):524–527
6. Sun XM, He MC (2005) Numerical simulation research on coupling support of roadway in soft rock at depth. J China Univ Min Technol 34(3):166–169
7. Mingzhong G (2007) Stiffness analysis of bolt and anchor cable supporting for coal roadway. J Min Saf Eng 24(4):396–400
8. Zhibiao G, Qian L, Jiong W (2009) Coupled bolt-mesh-anchor-tiusssupporting technology and its Engineering application to deep soft rocks roadway. Chin J Rock Mech Eng 28(2):3914–3919

Fast Distributed BFS Solution
for Edge-Disjoint Paths

Huiling Wu

Abstract We propose an improved synchronous distributed message-based breadth-first search (BFS) solution, to identify maximum cardinality sets of edge-disjoint paths, between a source node and a target node in a digraph. Previously, we presented a BFS based algorithm, NW-BFS-Edge (Nicolescu R, Wu H (2011) BFS solution for disjoint paths in P systems. In: Calude C, Kari J, Petre I, Rozenberg G (eds.) Unconventional computation, lecture notes in computer science, vol. 6714. Springer Berlin, pp 164–176), Nicolescu R, Wu H (2012) New solutions for disjoint paths in P systems. Nat Comput 11:637–651, which is a distributed version of the Edmonds-Karp algorithm Edmonds J, Karp RM (1972) Theoretical improvements in algorithmic efficiency for network flow problems. J ACM 19(2):248–264. Here, we propose a faster solution, Fast-NW-BFS-Edge, which improves NW-BFS-Edge by detecting and discarding "dead" cells in the first search round. On a set of randomly generated single-source directed acyclic graphs (dags), Fast-NW-BFS-Edge is 8.2 % faster than NW-BFS-Edge. This improvement has been inspired and guided by a P system modelling exercise, but is suitable for any distributed implementation.

Keywords Breadth-first search · Distributed · Edge-disjoint paths · Edmonds-Karp · Message-based · Network flow · P systems

1 Introduction

A P system is a parallel and distributed computational model inspired by the structure and interactions of living cells, introduced by Păun [1, 2]. A P system is specified by its membrane structure, symbols and rules. The structure is a digraph

H. Wu (✉)
Department of Computer Science, University of Auckland, Private Bag 92019 Auckland, New Zealand
e-mail: hwu065@aucklanduni.ac.nz

or a more specialised version, such as a directed acyclic graph (dag) or a tree. Each cell transforms its content symbols and sends messages to neighbours by applying rules. Rules of the same cell can be applied in parallel (where possible) and all cells work in parallel, traditionally in the synchronous mode.

The edge-disjoint paths problem finds a maximum cardinality set of edge-disjoint paths between a source node and a target node in a *digraph*. The classical algorithm transforms this problem to a maximum flow problem, solved by assigning unit capacity to each arc [3].

All algorithms discussed in this paper are distributed, totally message-based (no shared memory) and work synchronously: for brevity, we call them *distributed*. Dinneen et al. proposed the first depth-first search (DFS) based P system solution [4], as a distributed version of Ford-Fulkerson algorithm [5]. Our previous breadth-first search (BFS) based P system solution, NW-BFS-Edge, implements Edmonds-Karp algorithm [3] in a distributed way and improves its runtime, using a branch-cut technique [6, 2]. In this paper, our Fast-NW-BFS-Edge improves NW-BFS-Edge, by discarding "dead" cells, which are zero-out degree cells and the cells, of which all children are discarded. This pruning is performed in the first search round, parallel with the main search, without affecting the performance.

Our improvement is inspired by a P system modelling exercise, but is suitable for any distributed implementation. We assess the merits of P systems as directly executable formal specification of synchronous distributed algorithms and use them to benchmark our improved algorithm against our previous proposal.

2 Edge-Disjoint Paths in Digraphs

We consider a digraph, $G = (V, E)$, where V is a set of nodes and E is a set of arcs. For consistency with P system terminologies, nodes are also called cells. A set of paths are edge-disjoint if they have no common arc. Given a source node, $s \in V$, and a target node, $t \in V$, the edge-disjoint problem finds a maximum cardinality set of edge-disjoint *s-to-t* paths. This problem can be transformed to a maximum flow problem, by assigning unit capacity to each arc [3].

Given a set of edge-disjoint paths P, we defined $\overline{P}$ as the set of *path arcs*, (σ_i, σ_j), which appear in paths P, and $\overline{P}^{-1} = \left\{ (\sigma_j, \sigma_i) \mid ((\sigma_i, \sigma_j) \in \bar{P} \right\}$. The residual digraph $G_P = (V, E_P)$, where $Ep = \left(E \backslash \overline{P} \right) \cup \left(\overline{P}^{-1} \right)$. Briefly, the residual digraph is constructed by reversing path arcs. An augmenting path, α, is an *s-to-t* path in G_P. Augmenting paths are used to extend the set of edge-disjoint paths. If an augmenting arc reverses an existing path arc, then these two arcs "cancel" each other and are discarded. The remaining path fragments are relinked to construct an extended set of edge-disjoint paths. This process is repeated, starting with the new set of disjoint paths, until no more augmenting paths are found.
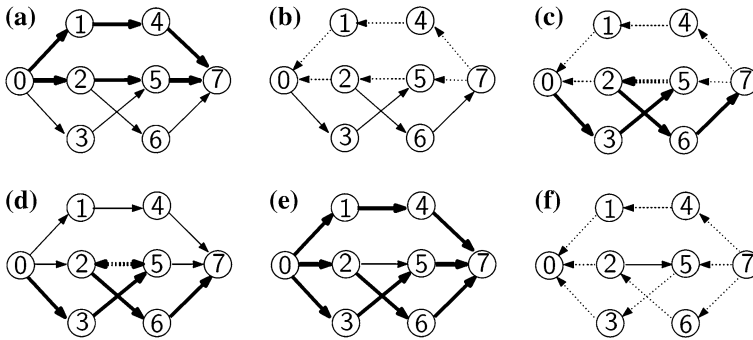
Fig. 1 Finding an augmenting path in a residual digraph. *Thin arcs* original arcs; *thick arcs* disjoint or augmenting path arcs; *dotted arcs* reversed path arcs

In Fig. 1, (a) shows the initial digraph, G, with two edge-disjoint paths, $P = \{\sigma_0.\sigma_1.\sigma_4.\sigma_7, \ \sigma_0.\sigma_2.\sigma_5.\sigma_7\}$; (b) shows the residual digraph, G_P, formed by reversing edge-disjoint path arcs; (c) shows an augmenting path, $\alpha = \sigma_0.\sigma_3.\sigma_5.\sigma_2.\sigma_6.\sigma_7$, which uses a reverse path arc, (σ_5, σ_2); (d) discards the cancelling arcs, (σ_2, σ_5) and (σ_5, σ_2); (e) relinks the remaining path fragments, $\sigma_0.\sigma_1.\sigma_4.\sigma_7, \ \sigma_0.\sigma_2, \ \sigma_5.\sigma_7, \ \sigma_0.\sigma_3.\sigma_5$ and $\sigma_2.\sigma_6.\sigma_7$, resulting in an incremented set of three edge-disjoint paths, $P' = \{\sigma_0.\sigma_1.\sigma_4.\sigma_7, \ \sigma_0.\sigma_2.\sigma_6.\ \sigma_7, \ \sigma_0.\sigma_3.\sigma_5.\sigma_7\}$ and (f) shows the new residual digraph, $G_{P'}$.

Augmenting paths can be searched using DFS [5] or BFS [7]. Conceptually, DFS explores as far as possible along a single branch before backtracking, while BFS explores as many branches as possible concurrently. A *search path* starts from the source and tries to reach the target. A *successful* search path becomes a new augmenting path and is used to increase the set of edge-disjoint paths: while conceptually a distinct operation, the new edge-disjoint paths are typically formed while the successful search path returns on its steps, back to the source.

When several search paths arrive, simultaneously or sequentially, the target decides which of them can succeed, using a *branch-cut* technique [6]. The branch ID of a search path is the cell ID of its first cell succeeding the source. Search paths sharing the same branch ID are incompatible; only one of them can succeed.

3 Our Fast BFS-Based Algorithm

Based on NW-BFS-Edge, our Fast-NW-BFS-Edge also works in successive search rounds. Round r starts with a set of edge-disjoint paths, which is empty at the start of the first search round. The source cell, σ_s, starts a search wave, to find new augmenting paths. Current "frontier" cells send out visit tokens to all its children. The *unvisited* cells, which receive and accept these visit tokens, set themselves as *temporarily visited* and become the new frontier.

The advancing wave periodically sends progress indicators back to the source: (a) extending notifications (at least one search path is still extending) and (b) path confirmations (at least one search path was successful, i.e. a new augmenting path was found). While moving towards the source, each path confirmation reshapes the existing paths and the newly found augmenting path, by discarding cancelling arcs and relinking the rest, building a larger set of edge-disjoint paths.

If no progress indicator arrives in the expected time, σ_s assumes that the round ends. If at least one search path was successful, σ_s initiates a global broadcast, to reset all temporarily visited cells to *unvisited*, which runs in parallel with the new search, without extra steps. Otherwise, σ_s initiates a global broadcast, to inform all cells that the algorithm has terminated.

Fast-NW-BFS-Edge improves NW-BFS-Edge, by discarding zero-out degree cells and the cells, of which all children are discarded, in round 1. If a frontier cell has no children, it discards itself, becoming *permanently visited*. Once discarded, a cell broadcasts permanently visited notifications to all its parents. When a previous frontier cell receives permanently visited notifications from all its children, it discards itself. This discarding process runs in parallel with the main search, along the extending notification of its search path, without extra steps.

Figure 2 shows how our optimisation works, in round 1. Subfigure (a) shows an original digraph. In (b), the search wave advances one step; σ_1, σ_2 and σ_3 become the frontier, sending back extending notifications, r's. In (c), the search wave advances one step; σ_4 and σ_5 become the frontier, sending back extending notifications, r's, and one search path reaches the target, σ_7, sending back a path confirmation, a_7. The frontier cell, σ_5, has no children, so it discards itself. In (d), the wave advances one step; σ_6 becomes the frontier, sending back an extending notification to σ_4. The frontier cell, σ_6, has no children, so it discards itself. In (e), the extending notification is relayed by σ_4–σ_1. Because all children of σ_4 are discarded, σ_4 discards itself. In (f), the extending notification is relayed by σ_1–σ_0. Because all children of σ_1 and σ_2 are discarded, σ_1 and σ_2 discard themselves. This discarding process runs in parallel with the main search, along the extending
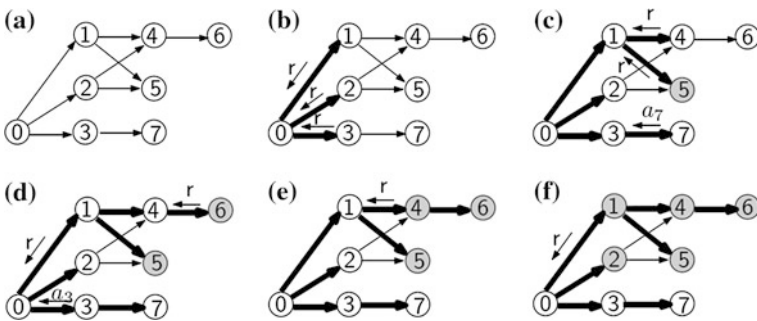
Fig. 2 *Thin arcs* original arcs; thick arcs: search path arcs; *gray cells* are discarded by Fast-NW-BFS-Edge; r is an extending notification; a_j is a path confirmation from cell σ_j

notification of its search path. In round 2, the source cell does not have any child to explore, so the algorithm terminates.

For NW-BFS-Edge, the round 1 search wave follows the same route as in Fast-NW-BFS-Edge, without our pruning optimisation. In round 2, because σ_1, σ_2, σ_4, σ_5 and σ_6 are not discarded (in round 1), the search needlessly probes them again. No search path succeeds, then the algorithm terminates. Both algorithms terminate in round 2, but Fast-NW-BFS-Edge is much faster because it does not explore any cell in round 2.

4 P Systems

We use a refined version of the *simple P module* [8], extended with generic rules using complex symbols [9] and matrix organised rules (here).

A simple P module is a system $\Pi = (V, E, Q, O, R)$, where V is a finite set of cells; E is a set of structural digraph arcs between cells, functioning as *duplex* channels; Q is a finite set of states; O is a finite non-empty alphabet of symbols; and R is a finite set of multiset rewriting rules (in a matrix format).

The general form of a rule in R is: $S\,x \to S'\,x'\,(y)\,\beta\gamma\ldots \mid z\,\neg z'$, where: $S, S' \in Q$, $x, x', y, z, z' \in O^*$ and ellipses (…) indicate possible repetitions of the last parenthesised item; S is known as the starting state and S' as the target state. This rule transforms state S to S'. It consumes multiset x and produces multiset x' ("here" multiset) in the same cell; in our matrix inspired formalism, x' becomes immediately available to other rules subsequently applied in the same step. Multiset y is a message queued and sent, at the end of the current step, as indicated by the transfer operator $\beta\gamma$. β's arrow indicates the transfer direction: ↑—to parents; ↓—to children; ↕—in both directions. γ indicates the distribution form: ∀—a broadcast; a structural neighbour, $\sigma_i \in V$—a unicast (to this neighbour). Multiset z is a promoter and z' is an inhibitor, which enables and disables the rule, respectively, without being consumed [2]. Operator α describes the rewriting mode. In the minimal mode, an applicable rule is applied once. In the maximal mode, an applicable rule is applied as many times as possible.

Matrix structured rulesets: Matrix structured rulesets are inspired by matrix grammars with appearance checking [10]. Ruleset R is organised as a matrix, i.e. a list of prioritised vectors: $R = (R_1, \ldots, R_m)$; all rules in a vector share the same starting state. Each vector R_i is a sequence of prioritised rules, $R_i = (R_{i,1}, \ldots, R_{i,mi})$.

The matrix semantics combines a strong priority for vectors and a weak-like priority for rules inside a vector. If a higher priority vector is applicable (at least one of its rules is applicable), then it is applied and all lower priority vectors are ignored; else, the next priority vector is considered. In a given vector, if a higher priority rule is applicable, then it is applied and the next priority rule is considered; else, the next priority rule is considered. After a rule is applied, "here" symbols become immediately available; outgoing messages are queued and sent at the end

of the step. This is the difference with the classical framework, where "here" symbols do not become available until the end of the step.

Complex symbols: We enhance our initial vocabulary, by recursive composition of elementary symbols into complex symbols, which are compound terms of the form: $t(i, \ldots)$, where (1) t is an elementary symbol representing the functor; (2) i can be (a) an elementary symbol, (b) another complex symbol, (c) a free variable (open to be bound, according to the cell's current configuration), (d) a multiset of elementary and complex symbols and free variables.

We often abbreviate complex symbols by using subscripts for term arguments. For example: $a(i) = a_i$, $e(2) = e_2$, $f(k, j) = f_{k, j}$, where a, e, f are elementary symbols and i, j, k are free variables. We assume that each cell σ_i is "blessed" with a unique *cell ID* symbol, $\iota(i)$, abbreviated as ι_i, which is exclusively used as a *promoter*.

Generic rules: To process complex symbols, we use high-level generic rules, which are instantiated using free variable matching [11]. A generic rule is identified by a combined *instantiation*. *Rewriting* mode, where (1) *instantiation* $\in$ {min, max} and (2) *rewriting* $\in$ {min, max}. The instantiation mode indicates how many instance rules are conceptually generated: (a) min indicates that the generic rule is nondeterministically instantiated only once, if possible; (b) max indicates that the generic rule is instantiated as many times as possible, without superfluous instances. The rewriting mode indicates how each instantiated rule is applied (as in the classical framework). For a detailed example, we refer readers to [9].

5 P System Specification

The input digraph is given by the P system structure and the system is fully distributed, i.e. there is no central node and only local channels (between structural neighbours) are allowed. Cells start without any topology awareness: they do not know the identities of their neighbours, not even their numbers.

P System Specification 1: Fast-NW-BFS-Edge

Input: All cells start with the same set of rules and without any topology awareness. All cells start in the same initial state, S_0. Initially, each cell, σ_i, contains an immutable cell ID symbol, ι_i. Additionally, the source cell, σ_s, and the target cell, σ_t, are decorated with symbols, s and t, respectively.

Output: All cells end in the same final state, S_{60}. On completion, all cells are empty, except that: (1) The source cell, σ_s, and the target cell, σ_t, are still decorated with symbols, s and t, respectively; (2) The cells on edge-disjoint paths contain path link symbols, for path predecessors, d'_j, and path successors, d''_k.

The full ruleset is detailed in our report. As an example, Table 1 shows the partial traces of the evolution of P system Specification 1, for Fig. 2.

Table 1 Partial traces of the evolution of P system specification 1, for Fig. 2

#	Cell	σ_0	σ_1	σ_2	σ_3
0	Init.	$S_0\ \iota_0\ s$	$S_0\ \iota_1$	$S_0\ \iota_2$	$S_0\ \iota_3$
1	Pre.	$S_{10}\ s\ n''_1\ n''_2\ n''_3$	$S_{20}\ n'_0\ n''_4\ n''_5$	$S_{20}\ n'_0\ n''_4\ n''_5$	$S_{20}\ n'_0\ n''_7$
2	(a)	$S_{10}\ s\ \ldots$	$S_{20}\ f_1\ldots$	$S_{20}\ f_1\ldots$	$S_{20}\ f_1\ldots$
3	(b)	$S_{10}\ s\ r^3\ldots$	$S_{20}\ v\ s'_0\ o_4\ o_5\ldots$	$S_{20}\ v\ s'_0\ o_4\ o_5\ldots$	$S_{20}\ v\ s'_0\ o_7\ldots$
4	(c)	$S_{10}\ s\ldots$	$S_{20}\ r^2\ v\ s'_0\ o_4\ o_5\ w_5$	$S_{20}\ v\ s'_0\ o_4\ o_5\ w_5\ldots$	$S_{20}\ v\ s'_0\ o_7\ a_7\ldots$
5	(d)	$S_{10}\ s\ a_4\ r\ldots$	$S_{20}\ v\ s'_0\ o_4\ u_5\ldots$	$S_{20}\ v\ s'_0\ u_5\ o_4\ldots$	$S_{20}\ v\ d'_0\ d''_7\ o_7\ldots$
6	(e)	$S_{10}\ s\ d''_4\ldots$	$S_{20}\ r\ v\ s'_0\ u_5\ o_4\ w_4$	$S_{20}\ v\ s'_0\ u_5\ o_4\ w_4\ldots$	$S_{20}\ v\ d'_0\ d''_7\ o_7\ldots$
7	(f)	$S_{10}\ s\ r\ d''_4\ w_2\ w_3\ldots$	$S_{20}\ w\ v\ s'_0\ u_4\ u_5\ldots$	$S_{20}\ w\ v\ s'_0\ u_4\ u_5\ldots$	$S_{20}\ v\ d'_0\ d''_7\ o_7\ldots$
8	Final	$S_{60}\ \iota_0\ s\ d''_4$	$S_{60}\ \iota_1$	$S_{60}\ \iota_2$	$S_{60}\ \iota_3\ d'_0\ d''_7$

#	Cell	σ_4	σ_5	σ_6	σ_7
0	Init.	$S_0\ \iota_4$	$S_0\ \iota_5$	$S_0\ \iota_6$	$S_0\ \iota_7\ t$
1	Pre.	$S_{20}\ n'_1\ n'_2\ n''_6$	$S_{20}\ n'_1\ n'_2$	$S_{20}\ n'_4$	$S_{30}\ t\ n'_3$
2	(a)	$S_{20}\ldots$	$S_{20}\ldots$	$S_{20}\ldots$	$S_{30}\ t\ldots$
3	(b)	$S_{20}\ f_{1.1}\ f_{2.2}\ldots$	$S_{20}\ f_{1.1}\ f_{2.2}\ldots$	$S_{20}\ldots$	$S_{30}\ t\ f_{3.3}\ldots$
4	(c)	$S_{20}\ v\ s'_1\ o_7\ldots$	$S_{20}\ w\ v\ s'_1\ldots$	$S_{20}\ f_{1.4}\ldots$	$S_{30}\ t\ d'_3\ e_3\ldots$
5	(d)	$S_{20}\ r\ v\ s'_1\ o_7\ w_7\ldots$	$S_{20}\ w\ v\ s'_1\ldots$	$S_{20}\ w\ v\ s'_4\ldots$	$S_{30}\ t\ d'_3\ e_3\ldots$
6	(e)	$S_{20}\ w\ v\ s'_1\ u_7\ldots$	$S_{20}\ w\ v\ s'_1\ldots$	$S_{20}\ w\ v\ s'_4\ldots$	$S_{30}\ t\ d'_3\ e_3\ldots$
7	(f)	$S_{20}\ w\ v\ s'_1\ u_7\ldots$	$S_{20}\ w\ v\ s'_1\ldots$	$S_{20}\ w\ v\ s'_4\ldots$	$S_{30}\ t\ d'_3\ e_3\ldots$
8	Final	$S_{60}\ \iota_4$	$S_{60}\ \iota_5$	$S_{60}\ \iota_6$	$S_{60}\ \iota_7\ t\ d'_3$

Row 0 indicates the initial contents (input). Row 1 indicates the contents after the preliminary phase, local topology discovery: n'_j points to structural parents; n''_k points to structural children. Rows 2–7 correspond to Fig. 2a–f: omitted symbols (…) are cell ID symbols and pointer symbols of structural parents and children; d'_j points to edge-disjoint path predecessors; d''_k points to edge-disjoint path successors; s'_j points to search path predecessors; f_j is the visit token from the source; $f_{j.k}$ is the visit token from σ_k of branch j; r is an extending notification; a_j is the path confirmation from σ_j; e_j records a branch ID j in the target; w_j is the permanently visited notification from σ_j; u_j records that σ_j is permanently visited; w indicates that the cell is permanently visited; v indicates that the cell is temporarily visited. Row 8 indicates the final contents (output).

6 Runtime Performance

Consider a digraph with n cells, where f is the maximum number of edge-disjoint paths. In both NW-BFS-Edge and Fast-NW-BFS-Edge, at least one augmenting path is found in each search round, so the number of search rounds is at most f. In each round, the search wave reaches the target in at most n steps. As earlier mentioned, all housekeeping operations are performed in parallel with the main search. Thus, Fast-NW-BFS-Edge runs in $O(nf)$ steps (same as NW-BFS-Edge).

Table 2 Performance on 30 random single-source dags with $n = 100$ nodes

m	250	500	1,000	1,500	200	2,500	3,000
Speed-up	7.3 %	6.5 %	5.4 %	6.9 %	7.7 %	11.6 %	11.7 %

However, this theoretical estimation does not fully count for the detection and discarding of "dead" cells. Therefore, we experimentally compare NW-BFS-Edge and Fast-NW-BFS-Edge, using their executable P specifications.

Although Fast-NW-BFS-Edge works for diagraphs, we feel that dags take more advantage of our optimisation. For Fig. 2, Fast-NW-BFS-Edge takes 23 steps, showing 25.8 % speed-up over NW-BFS-Edge, which takes 31 steps.

We further test two algorithms on a series of random single-source dags generated using NetworkX [12]. Each set has thirty single-source dags with $n = 100$ nodes and m arcs, where $m \in \{250, 500, 1,000, 1,500, 2,000, 2,500, 3,000\}$. Table 2 shows the average speed-up of Fast-NW-BFS-Edge over NW-BFS-Edge in each set. We also compute the global average speed-up over all these scenarios. Overall, Fast-NW-BFS-Edge is 8.2 % faster than NW-BFS-Edge.

7 Conclusion

We presented a fast distributed BFS-based algorithm, Fast-NW-BFS-Edge, for the edge-disjoint paths problem. Using a novel idea, Fast-NW-BFS-Edge discards "dead" cells detected in the first search round. We provided a directly executable formal P specification, which uses high-level generic rules organized in a matrix-like structure. The resulting P system has a reasonably fixed-size ruleset and achieves the same runtime complexity as the corresponding distributed algorithm.

Our Fast-NW-BFS-Edge algorithm works for digraphs and empirical results show that it achieves worthwhile performance improvement over Fast-NW-BFS-Edge on single-source dags, especially for high-density dags. On the other side, one can construct many sample scenarios where Fast-NW-BFS-Edge vastly outperforms NW-BFS-Edge.

Several interesting questions remain open. How does the speedup depend on the digraph density? Can these results be extrapolated to digraphs with different characteristics, such as size, average node degree, node degree distribution? Are there other possible optimisations? Are there well defined practical scenarios where one could recommend Fast-NW-BFS-Edge over NW-BFS-Edge?

References

1. Păun G (2000) Computing with membranes. J Comput Syst Sci 61(1):108–143
2. Păun G, Rozenberg G, Salomaa A (2010) The oxford handbook of membrane computing. Oxford University Press, New York
3. Cormen TH, Stein C, Rivest RL, Leiserson CE (2009) Introduction to algorithms (3rd edn). The MIT Press, Cambridge
4. Dinneen MJ, Kim YB, Nicolescu R (2010) Edge- and node-disjoint paths in P systems. Electron Proc Theor Comput Sci 40:121–141
5. Ford LR Jr, Fulkerson DR (1956) Maximal flow through a networkX. Can J Math 8:399–404
6. Nicolescu R, Wu H (2011) BFS solution for disjoint paths in P systems. In: Calude C, Kari J, Petre I, Rozenberg G (eds.) Unconventional computation, lecture notes in computer science, vol. 6714. Springer Berlin, pp 164–176
7. Edmonds J, Karp RM (1972) Theoretical improvements in algorithmic efficiency for network flow problems. J ACM 19(2):248–264
8. Dinneen MJ, Kim YB, Nicolescu R (2010) A faster P solution for the Byzantine agreement problem. In: Gheorghe M, Hinze T, Păun G (eds.) Conference on membrane computing. Lecture Notes in Computer Science, Vol. 6501. Springer, Berlin, pp 175–197
9. Nicolescu R (2012) Parallel and distributed algorithms in P systems. Lect Notes Comput Sci 7184:33–42
10. Freund R, Păun G (1995) A variant of team cooperation in grammar systems. J Univ Comput Sci 1(2):105–130
11. Bălănescu T, Nicolescu R, Wu H (2011) Asynchronous P systems. Int J Nat Comput Res 2(2):1–18
12. Hagberg AA, Schult DA, Swart PJ (2008) Exploring network structure, dynamics, and function using network. In: Varoquaux G, Vaught T, Millman J (eds.) 7th Python in Science Conference (SciPy), pp 11–15
13. Nicolescu R, Wu H (2012) New solutions for disjoint paths in P systems. Nat Comput 11:637–651

Stock Price Forecast Using Tree Augmented Naïve (TAN) Bayes

Yuanlu Liao, Xiangxiang Zeng, Tao Song and Lianfeng Zhang

Abstract The Tree Augmented Naïve (TAN) Bayes is one of Bayesian network algorithm and it can create a simple Bayesian network model efficiently. This paper describes a method using TAN to create a Bayesian network which can help to forecast the change rate of Chinese Company's stock price change per day. Firstly, we collect Wanke Company's stock price data and discrete these data with discrete algorithm. Secondly, we use the discrete data to construct the TAN network. Thirdly, we use TAN network to forecast the Wanke Company's stock price data with variable correlation coefficient and the root mean square error as estimators.

Y. Liao · X. Zeng (✉)
Department of Computer Science, Xiamen University, Xiamen 361005,
People's Republic of China
e-mail: xzeng@xmu.edu.cn

Y. Liao
e-mail: yuanluliao@Foxmail.com

T. Song
Department of Control Science and Engineering, Huazhong University of Science
and Technology, Wuhan 430074, Hubei, People's Republic of China
e-mail: songtao0608@hotmail.com

L. Zhang
Qingdao Branch, Naval Aeronautical Engineering Institute, Qingdao 266000,
Shandong, People's Republic of China
e-mail: zlf532@126.com

1 Introduction

Stock market forecast is the act of trying to determine the future value of a company stock or other financial instrument traded on a financial exchange. The successful prediction of a stock's future price could yield significant profit. There are many time-series algorithm have been applied in the stock price forecast [1, 2], have been applied in the stock price forecast. For example, Auto Regressive (AR) model, Moving Average (MA) model, Auto Regressive Moving Average (ARMA) model and Auto Regressive Conditional Heteroscedasticity (ARCH) model [3].

The disadvantage of above time-series algorithm is that they usually represent the error distribution according to the normal distribution which is inconsistent with recent study. The study shows that the distribution of the stock price fluctuation does not follow the normal distribution [4].This makes above time-series algorithm cannot forecast accurately. So some researchers presented the Bayesian network method [5] using the K2 algorithm [6, 7], with K2 Metric [6–8], as the estimator of the network and applied the method to forecast Japan Company's stock price.

The Tree Augmented Naïve (TAN) Bayes [9] is one of Bayesian network algorithm and it can create a simple Bayesian network model efficiently. In this paper, we introduce TAN Bayes method and applied the method into the Chinese stock. Because Bayesian network can just deal with discrete value, this paper also introduce two different discrete algorithms, uniform and ward, to discrete relevant data.

The paper is organized as follows. In the next section, we will introduce discrete algorithm with the notation that we will use in that section. In Sect. 3, we show an example of applying TAN method into Chinese stock. We conclusion and show the future work in Sect. 4 and describe the acknowledgment about this paper in "Acknowledgment".

2 Notation and Discrete Algorithm

2.1 Notations

The change rate of stock price per day defines as follows.

$$r_t = (\ln(p_t) - \ln(p_{t-1})) * 100 \tag{1}$$

where the notation p_t denotes the closing stock price at date t.

We calculate ach r_t and sort them in ascending order and divide these data into L clusters, show as follow.

$$\{C_1, \ C_2, \ \ldots, C_l\}$$

where the notation C_l (l = 1,2,…,L) denote the cluster and we use the notation c_l(l = 1,2,…,L) denote the cluster C_l's center.

2.2 Uniform Algorithm

The uniform algorithm divides the ordered r_t into L clusters making sure each cluster have as possible as same number of each samples.

2.3 Ward Algorithm

The ward algorithm divides the ordered r_t into L clusters making sure the sum of Euclid distances from each sample to its cluster's center is minimum.

$$D(C_i, C_j) = E(C_i \cup C_j) - E(C_i) - E(C_j) \tag{2}$$

$$E(C_i) = \sum_{z \in C_i} d(z, c_i)^2 \tag{3}$$

where the notation z denotes all samples in the cluster C_i and $d(z, c_i)$ denotes the Euclid distance between c_i and z. E denotes Euclid distance.

3 Stock Price Forecast Test

3.1 Notations

In this paper, we use the variable correlation coefficient (CC) and the root mean square error (RMSE) to estimate the forecast. They are summarized as follows.

$$CC = \frac{\sum_{t-1}^{n}(r_t - \bar{r})(r_t' - \bar{r}')}{\sqrt{\sum_{t=1}^{n}(r_t - \bar{r})^2} * \sqrt{\sum_{t=1}^{n}(r_t' - \bar{r}')^2}}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(r_t - r_t')^2}$$

$$\bar{r} = \frac{1}{n}\sum_{t=1}^{n} r_t \quad \bar{r}' = \frac{1}{n}\sum_{t=1}^{n} r_t'$$

where the notation r_t denote the actual stock prices and r_t' denote the predicted stock prices. n describe the total number of the change rate of stock price per day.

3.2 Wanke Company's Stock Price Forecast

To evaluate the performance of the TAN network, we use the Wanke Company's stock price to have a test. We construct a TAN Bayesian network and use this network to predict the change rate of stock price. At last we use variable CC and RMSE as the estimators to estimate this TAN method.

In discrete algorithm, we divide r_t into L = 6 clusters. The result information of the cluster created by uniform algorithm is show in Table 1 and the TAN network is showed in Fig. 1. The result information of the cluster created by ward algorithm is show in Table 2 and the TAN network is showed in Fig. 2. Where the notation $(C_l)_{min}$ and $(C_l)_{max}$ in tables are denote the minimum and maximum values of the samples in the cluster C_l.

The truth distribution of the change rate of stock price per day (truth r_t) is shown in Fig. 3. The distribution of the change rate of stock price per day (predict r_t) using uniform algorithm is shown in Fig. 4. The distribution of the change rate of stock price per day (predict r_t) using ward algorithm is shown in Fig. 5.

Table 1 Classified section of Wanke Company's r_t using uniform algorithm

Cluster	$(C_l)_{min}, (C_l)_{max}$	Number of sample	C_l (%)
C1	[−70.99, −2.04 %]	878	−4.62
C2	[−2.03, −0.87 %]	875	−1.04
C3	[−0.86, −0.03 %]	815	−0.46
C4	[0.00, 0.73 %]	892	0.30
C5	[0.74, 2.19 %]	896	1.34
C6	[2.20, 39.57 %]	896	4.61

Fig. 1 Bayesian network for Wanke using uniform algorithm
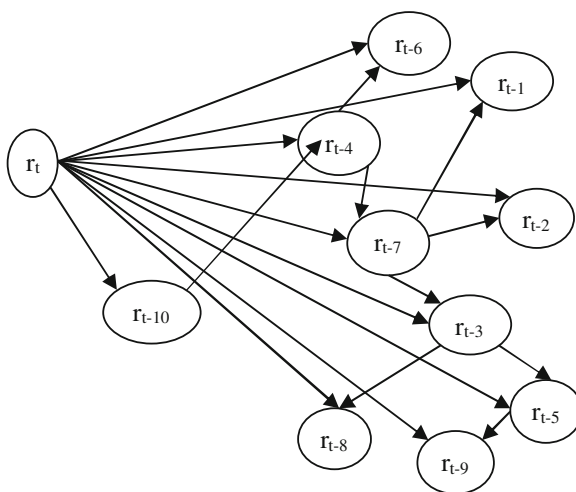
Table 2 Classified section of Wanke Company's r_t using ward algorithm

Cluster	$(C_l)_{min}$, $(C_l)_{max}$	Number of sample	C_l (%)
C1	[−70.99, −3.22 %]	497	−6.17
C2	[−3.21, −0.39 %]	1758	−1.39
C3	[−0.38, 0.86 %]	1331	0.21
C4	[0.87, 2.13 %]	750	1.39
C5	[2.14, 6.33 %]	762	3.47
C6	[6.37, 39.57 %]	154	9.75

Fig. 2 Bayesian network for Wanke using ward algorithm
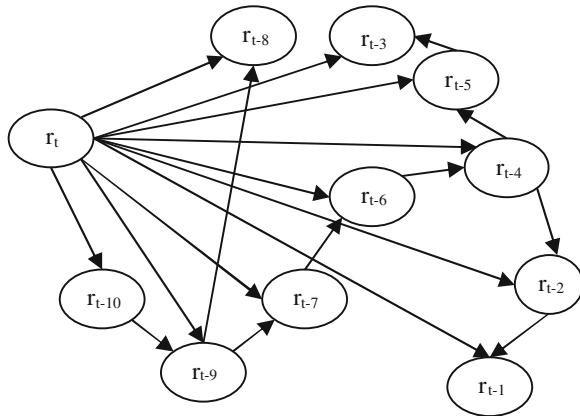
Fig. 3 Truth distribution of change rate of stock price per day
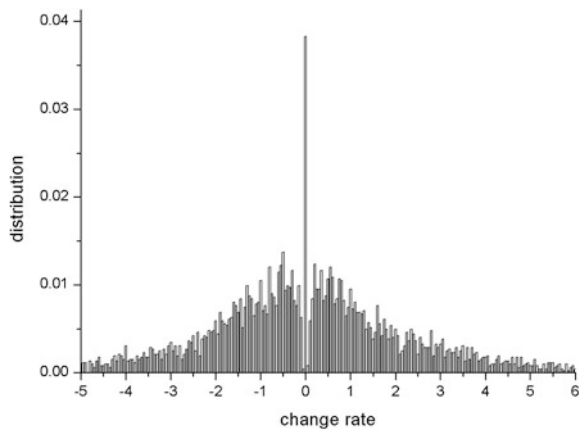
3.3 CC and RMSE

We can use the TAN Bayesian network to predict Wanke Company's stock price. This paper uses the CC and RMSE as the estimator. In discrete algorithm, we divide r_t into L = 6 clusters.

Fig. 4 Distribution after use uniform
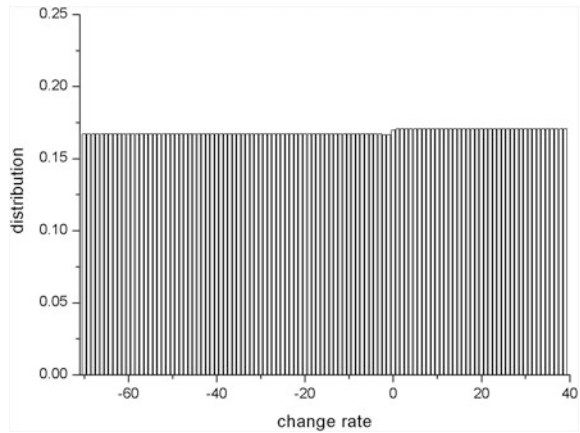
Fig. 5 Distribution after use ward
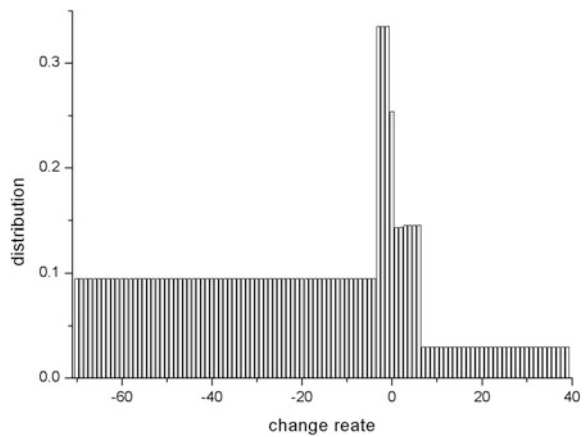
When we use the uniform algorithm to discrete, we find that RMSE $= 3.719563$ and CC $= 0.182219$. If we use the ward algorithm, we find that RMSE $= 3.717633$ and CC $= 0.16866$.

4 Conclusion and Future Work

This paper uses the TAN Bayesian network to forecast the change rate of stock price per day. Firstly, we calculate r_t and we use the uniform algorithm or ward algorithm to discrete r_t. Secondly, we use the discrete r_t to construct the TAN Bayesian network and finish the TAN Bayesian network's parameter learning. At

last, we use the Wanke Company's stock price to test this method with variable CC and RMSE as the estimator.

There are many other Bayesian network modes, so it will be interesting to use other Bayesian network methods in stock price prediction. On the other hand, more experiment can be done to test TAN network with other company's stock price as data.

References

1. Box G, Jenkins GM, Reinsel GC (1994) Time series analysis: forecasting and control. Prentice-Hall
2. Brockwell P, Davis R (2009) Time series: theory and methods. Springer, Berlin
3. Engle RF, Ng VK (1993) Measuring and testing the impact of news on volatility. J Financ 48(5):1749–1778
4. Takayasu H (2006) Practical fruits of econophysics. Springer, Berlin
5. Zuo Y, Kita E (2011) Stock price forecast using Bayesian network. Expert Syst Appl 39(8):6729–6737
6. Ben-Gal I (2007) Bayesian network. In Ruggeri F, Ruggeri RS, Kennett, Faltin FW (eds.), Encyclopedia of statistics in quality and reliability, Wiley, London
7. Pearl J, Russell S (2002) Bayesian network. In Arbib MA (eds.), Handbook of brain theory and neural network, MIT press, Cambridge, pp 157–160
8. Heckerman D, Geiger D, Chickering D (1995) Learning Bayesian networks: the combination of knowledge and statistical data. Mach Learn 20:197–243
9. Chow C, Lui C (1968) Approximating discrete probability distributions with dependence trees. IEEE Trans Info Theor 14:462–467

A Hybrid Membrane Computing and Honey Bee Mating Algorithm as an Intelligent Algorithm for Channel Assignment Problem

Maroosi Ali and Ravie Chandren Muniyandi

Abstract Membrane computing is a model of computation inspired by the structure and functioning of cells as living organisms. Membrane computing naturally has parallel structure. Also, it uses communication rules to exchange information between membranes. This paper first proposes Hybrid Honey Bee Mating (HHBM) then uses parallelism advantage of membrane to parallelize and divide the HHBM algorithm as an evolutionary algorithm to different membranes (parts). These membranes can be executed in parallel way on different cores or CPUs. Simulation shows that when number of membrane increases performance of this algorithm increases.

Keywords Membrane computing · Intelligent algorithm · Channel assignment

1 Introduction

The theoretical model used in membrane computing is known as a P System introduced by Gheorghe Paun [1]. Membrane computing mostly is used to simulate biological cases [1]. Although bio inspired, membrane computing is also being applied to problems beyond biology, including modeling, economics, databases, networks, hard problem and so on. One common goal of membrane computing is to be able to solve computationally hard problems, usually in polynomial or even linear time. In this paper we use from power of membrane

M. Ali (✉) · R. C. Muniyandi
Research Center for Software Technology and Management, Faculty of Technology and Information Science, National University of Malaysia, 43600 Bangi, Selangor, Malaysia
e-mail: ali.maroosi@gmail.com

R. C. Muniyandi
e-mail: ravie@ftsm.ukm.my

computing and heuristic algorithm [2] to solve channel assignment problem as a NP hard problem.

The interference constraints in a cell network are usually described by symmetric matrix called compatibility matrix C [3]. Where c_{ij} denotes the minimum required channel separation between a channel assigned to a call in cell i and a channel assigned to another call in cell j. If $c_{ij} = c_{ij}$ there is cosite constraint (CCC) $c_{ij} = 0$ there is no constraint in channel reuse. A demand vector $m = (m_i) (1 \leq i \leq n)$, where m_i indicates the number of channels required for cell i. In this paper, number of blocked calls b should be force to zero (b = 0) by appropriate channel assignment algorithm. This paper proposed new hybrid channel assignment that is combination honey bee mating Algorithm [2] and a local search algorithm and characteristic of membrane computing [1].

2 The Proposed Algorithms

2.1 Proposed Hybrid Honey Bee Mating Algorithm

Proposed algorithm consists of the following three channel assignment stages (The original HBM algorithm explained with more detail in [2]):

Stage 1: Regular interval assignment

The first stage assigns channels to calls in the cell that determines the lower bound on the number of channels M. This is typically the cell that has the largest call demand.

Stage 2: Greedy assignment

Node-degree ordering: Let the degree d_i be a measure of the difficulty of assigning a channel to cell i. It is defined as follows:

$$d_i = \sum_{1}^{N} m_j c_{ij} - c_{ii} \text{ if } m_j \neq 0 \text{ else } d_i = 0. \tag{1}$$

Frequency exhaustive strategy and requirement exhaustive strategy: Once the cells have been sorted based on node degree FEA starts with the first cell in the ordered list. Pseudo code of frequency exhaustive assignment strategy as follows:

```
Loop 1: for each call i in the call list.
      Loop 2: for each channel j within the lower bound
              IF i is not assigned AND j can be assigned to i
                      Assign j to i
                      Break Loop2
              End If
      End Loop 2
End Loop 1
```

Greedy assignment stage: The greedy assignment stage assigns channels to cells that form the greedy region. The greedy region is regarded as the region of cells with the highest assignment difficulties. The procedure of this stage is as follows:

1. Initialize the greedy region to include the cell with the largest node-degree and all six of its adjacent cells.
2. Make a list of calls from the greedy region in the ascending order of the cell index number.
3. Apply FEA to the channels in the call list.

If every call in the list is assigned a channel, then terminate the procedure as successful and proceed to the honey-bee mating algorithm stage. Else, expand the greedy region by including all cells adjacent to the existing greedy region. Then, go back to Step 3. If the greedy region assignment fails, go to Stage 3.

Stage 3: Application of Hybrid honey-bee mating optimization algorithm. Once the first two stages of our algorithm are completed, calls in the remaining cells are selected and then these steps executed.

Step 1: String representation: Each population element, or a so-called solution point x^t ($t = 1, \ldots, K$) represents a vector of the randomly ordered cell assignment of the remaining cells. Then the fitness of each individual is evaluated. FEA is used to assign channels to the call list. The total number of blocked calls (b) for each list is used as the fitness.

Step 2: Define the model inputs parameters.

Step 3: Random generation of a set of initial solutions: This paper developed a different encoding according to the order of cells instead of call list [4] for assigning channels. If a_{ik}, the kth call in the ith cell and summation of total number of call in all cells named M. Then call list represented as $X = [a_{i_1 k_1}, a_{i_2 k_2}, \ldots, a_{i_M k_M}]$ where a_{ik}, is the kth call in the ith cell. Therefore this paper represent these call lists according to cell number i.e. $X(new) = [i_1, i_2, \ldots, i_M]$ where i_j is the cell number that call in jth order in list belong to this cell in list. If encoding represented according calls there are $M!$ different call list but if represent according this paper encoding there are $M!/(m_1! \times m_2! \times \ldots \times m_{n-1}! \times m_n!)$ (M is total calls, ! is factorial operator, m_i represent number of calls in cell i.) different list for representation without loses any information.

Step 4: Selection of queen: After generation of a set of initial solutions mentioned in the previous stage, rank the initial solutions based on performance function and set the best one as queen.

Step 5: Mating flight: Use simulated annealing to select the set of solutions from the search space to make a mating pool for possible information exchange between the best present solution (queen) and the selected trial solutions.

Step 6: Breeding process: Generate new set of solutions by employing predefined crossover operators and heuristic functions between the present solutions and the trial solution according to their fitness values. In this study, select one parent cell lists List1 and Queen. Then randomly generate a binary. Afterwards, if the element of the binary vector at the corresponding position has a value of one, then
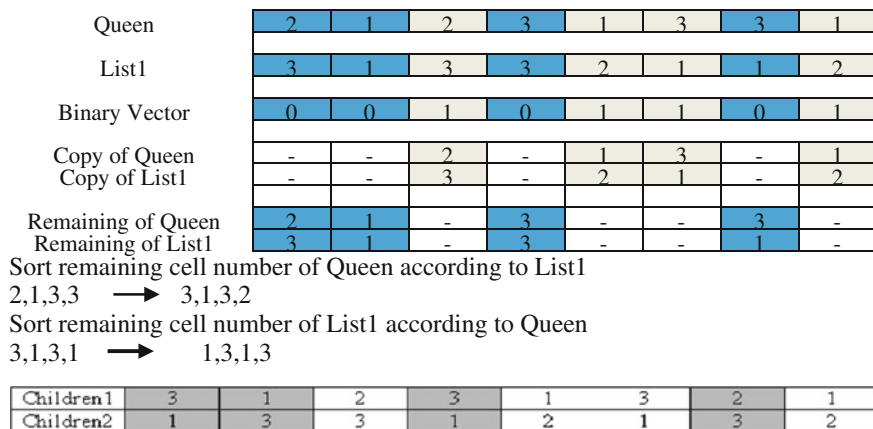
Queen	2	1	2	3	1	3	3	1
List1	3	1	3	3	2	1	1	2
Binary Vector	0	0	1	0	1	1	0	1
Copy of Queen	-	-	2	-	1	3	-	1
Copy of List1	-	-	3	-	2	1	-	2
Remaining of Queen	2	1	-	3	-	-	3	-
Remaining of List1	3	1	-	3	-	-	1	-

Sort remaining cell number of Queen according to List1
2,1,3,3 $\longrightarrow$ 3,1,3,2
Sort remaining cell number of List1 according to Queen
3,1,3,1 $\longrightarrow$ 1,3,1,3

Children1	3	1	2	3	1	3	2	1
Children2	1	3	3	1	2	1	3	2

Fig. 1 Crossover between queen and list 1 and generate children 1 and children 2

copy the elements of Queen to the same positions within the new list. The remaining elements of Queen are permutated so that they appear in the same order as in List1. Next, enter these resorted elements in the empty gaps of the new cell list. An example of crossover is illustrated in Fig. 1.

Step 7: feeding selected broods and queen with the royal jelly by workers: In this study, Children are mutated. Figure 2 shows the mutation procedure. If a random number from 0 to 1 for each gene is less than the mutation rate, then the gene is selected for mutation. The selected genes (3, 2, 1, and 3) are randomly rearranged during mutation.

Step 8: Local search (Like CAP3): CAP3 [5] exchanges the gene with the lowest channel number with a randomly selected gene. Since in this paper only channels below the lower bound frequency (in Table 2) assigned to call. So we determine cells with blocked call and exchange order of call within blocked cell with call in non- blocking cells during the FEA.

Step 9: If the new best solution is, better than the queen replace it with queen.

Step 10: Check the termination criteria: If the termination criteria satisfied finish the algorithm, else discard all previous trial solutions and generate new trial solutions. Then go to step 5 until all assigned iteration (mating flights) completed or convergence criteria met.

2.2 Proposed Membrane Computing Honey Bee Mating Algorithm

First individuals generate by m different membranes on m exist cores (CPUs) on computers. Then for certain times this individuals improved by HHBM algorithms
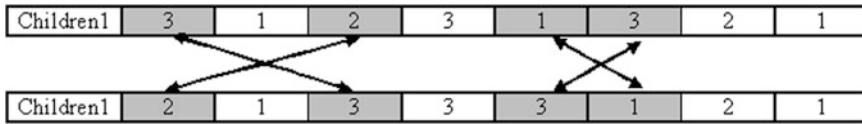
| Children1 | 3 | 1 | 2 | 3 | 1 | 3 | 2 | 1 |

| Children1 | 2 | 1 | 3 | 3 | 3 | 1 | 2 | 1 |

Fig. 2 Mutation of solutions

Table 1 Specifications for Sivarajan's benchmark problem

Number of instance	Compatibility matrix				Demand vector	
	Nc	Cij	ACC	Cii	Call case	Total calls
1	12	1	2	7	1	481
2	7	1	2	7	1	481
3	12	1	1	5	1	481
4	7	1	1	5	1	481
5	12	1	1	7	1	481
6	7	1	1	7	1	481
7	7	1	2	5	2	470
8	12	1	2	7	2	470
9	7	1	2	7	2	470
10	12	1	2	12	2	470

on different cores. After this step they communicate with each other as Fig. 3 and this repeated for several times.

In communication step all individuals in membranes gathered form cores to master core. According to the Tissue like P system rules [6] these individuals exchanged between membranes base on $(i, u/v, j)$ for $i, j \in \{1, 2, \ldots, m\}$, $i \neq j$, rules. This paper define $u = I^i_{l \times m + (j-i)+1}$ and $u = I^j_{l \times m + (m-(j-i))+1}$ that we assume number of individuals (n_i) in all m membranes are same i.e. $(n_1 = \ldots = n_i = \ldots = n_m)$, $j > i$, $l = 0, 1, \ldots, \frac{n_i}{m}$ and number of individuals (n_i) are multiple of number of membranes(m). By applying these rules membranes exchange information (In this paper information is individuals that contain information about solution of problem).

For example individuals for 3 membranes with 6 individuals inside each membrane $\{I^i_1, I^i_2, \ldots, I^i_6\}$; $i = 1, 2, 3$ after applying the rules are as Fig. 4a:

Fig. 3 Communication between four different membranes
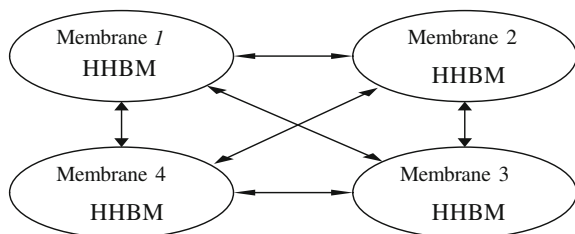
Results of these rules for each membrane similar to this that start from membrane i and first individual of this membrane selected as new first individual in membrane i then change the membrane as Fig. 4b and choose second individual of next membrane as new second individual in membrane i then again change membrane and select third individual in it as new third individual in membrane i and do this until all n_i new individuals are chosen for membrane i then repeat this for other membranes.
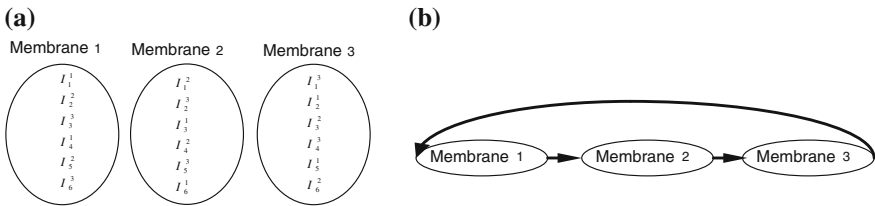
3 Simulations and Results

In this paper the Philadelphia problem is solved. In the compatibility matrix of Table 1, Nc is the cluster size for CCC. C_{ij} is the minimum channel distance between any pair of cells in the same cluster except for adjacent cells for CCC. Adjacent channel constraint (ACC) is the minimum channel distance between adjacent cells. C_{ii} is the minimum distance within the same cell for cosite constraint (CSC). In the demand vector, Call Cases 1 and 2 represent the corresponding channel requirements m_1 and m_2, respectively:

m1 = {8, 25, 8, 8, 8, 15, 18, 52, 77, 28, 13, 15, 31, 15, 36, 57, 28, 8, 10, 13, 8}

m2 = {5, 5, 5, 8, 12, 25, 30, 25, 30, 40, 40, 45, 20, 30, 25, 15, 15, 30, 20, 20, 25}

Simulation results shows that hybrid MCHBM algorithm can solve all instance in Sivarajan's benchmark [7].

In this paper proposed algorithm simulated on computer with Intel core i5 2.5 GHz (two cores) and 4 GB RAM with the parameter number drones 50, capacity of spermatheca 50, speed of Queen randomly $\in[0.5–1]$, minimum speed of Queen randomly $\in[0–1]$ speed reduction 0.98 that these parameter obtain experimentally. Table 2 illustrates the simulation results of different channel assignment algorithms. The Hybrid Membrane Computing Honey Bee Mating (HMCHBM) algorithm proposed in this paper successfully found the lower bound solution in all of 10 instances. From our simulations, for the easy cases, Number 1–6 and Number 8–10, we can find an optimal solution by using the first two stages of new algorithm and sometimes with just one or two generations of the third stage. For case 7,

Fig. 4 Exchanging the individuals (**a**) exchanged individuals between three membranes (**b**) Order of choosing individuals from membranes

Table 2 Simulation results for Sivarajan's and the new hybrid MCHBM algorithm

Number of instance	Lower bound	Sivarajan's	Hybrid MCHBM
1	533	536	533
2	533	533	533
3	381	381	381
4	381	381	381
5	533	533	533
6	533	533	533
7	258	269	258
8	309	310	309
9	309	310	309
10	529	529	529

usually require over a lot of generations of the third stage to achieve the optimum solution.

It takes time for master core to allocate or collect the data from other cores for the HMCHBM. This algorithm is so useful for problems that take several seconds to be solved because the overhead time for parallel run is neglectable respect to process time. For example, we execute both HHBM and HMCHBM for 200 s to solve channel assignment problem instance 7 the results of output are as follows:

	Percentage of successful assignment all calls in instance 7 in 200 s (%)
Proposed HHBM	77
Proposed HMCHBM run on two cores	100

For other instance since in less than 3 s solved it is better to use proposed HHBM instead of HMCHBM.

4 Conclusion

This paper presented a new algorithm to solve the NP-complete channel assignment problem in a cellular mobile communication system. This algorithm is hybrid of a HBM (as a global search) and the like CAP3 that is a local search algorithm and uses advantage of membrane computing. This algorithm is distributed like membrane computing in different membranes (parts) these membranes are executed over different cores (CPUs) and communicate and exchange information with each other to increase diversity of algorithm. In previous papers [8, 9] just use communication advantages of membrane computing to improve evolutionary algorithms. This paper uses both communication and parallel structure by defining appropriate algorithm that can be executed over different cores (CPUs) in parallel.

References

1. Paun Gh (2000) Computing with membranes. J Comput Syst Sci 61:108–143
2. Fathian M, Amiri B, Maroosi A (2007) Application of honey-bee mating optimization algorithm on clustering. Appl Math Comput 190:1502–1513
3. Abril J, Comellas F, Cortes A, Ozon J, Vaquer M (2000) A Multiagent system for frequency assignment in cellular radio networks. IEEE Trans Veh Technol 49:1558–1565
4. Beckmann D, Killat U (1999) A New strategy for the application of genetic algorithms to the channel-assignment problem. IEEE Trans Veh Technol 48:1261–1269
5. Wang W, Rushforth CK (1996) An adaptive local-search algorithm for the channel-assignment problem (CAP). IEEE Trans Veh Technol 45(3):456–466
6. Martin-Vide C, Paun Gh, Pazos J, Rodriguez-Paton A (2003) Tissue P systems. Theoret Comput Sci 296:295–326
7. Sivarajan KS, McEliece RJ, Ketchum JW (1989) Channel assignment in cellular radio. In: Proceedings of 39th IEEE vehicular technology conference, pp 846–850
8. Zhang G, Cheng J, Gheorghe M, Meng Q (2012) A hybrid approach based on differential evolution and tissue membrane systems for solving constrained manufacturing parameter optimization problems. J Appl Soft Comput, In press, http://dx.doi.org/10.1016/j.asoc.2012.05.032
9. Cheng JX, Zhang GX, Zeng XX (2011) A novel membrane algorithm based on differential evolution for numerical optimization. Int J Unconventional Comput 7:159–183

A Salt and Pepper Noise Image Filtering Method Using PCNN

Rencan Nie, Shaowen Yao, Dongming Zhou and Xiang Li

Abstract Based on Pulse Coupled Neural networks, an effective salt and pepper noise image filtering method is proposed. Similar groups of neurons burst synchronous pulses in a PCNN, thereby the noise pixels are detected. Considering a noise pixel has the most similarity with neighbor non-noise pixels, a filtering method called extended median filtering is put forward to filter the noises in an image. Simulation results show that the method proposed has excellent filtering performance for the noise images with different noise intensity, and has the obvious advantage compared with the median filter.

Keywords Salt and pepper noise · Filtering · Pulsed coupled neural network · Extended median filter

1 Introduction

Salt and pepper noise is a typical random noise, and it locates in the high or low end in the range of the image gray-scale, so the noises are black or white points in an image. As a result, the image features will be greatly damaged. Median filter (MF) [1] as a nonlinear filtering method is a typical method for salt and pepper noise. But the MF processes noise pixels and non-noise pixels uniformly without discrimination. This leads to the information damaged by non-noise pixels. On the other hand, the filtering performance by MF declines sharply with the increase of noise intensity. Thus, a filtering technique based on noises detection is widely applied in its improvement.

R. Nie (✉) · D. Zhou · X. Li
Information College, Yunnan University, Kunming 650091, China
e-mail: Rencan Nie@163.com; huomu_ren@163.com

S. Yao
Software College, Yunnan University, Kunming 650091, China

Pulse Coupled Neural Network (PCNN) is a new artificial neural network, which was obtained by modeling from the phenomena of bursting synchronous pulses in cat visual cortex [2]. It has been widely applied in the areas such as image segmentation [3], image fusion [4, 5], image de-noising [6, 7] and so on. For a salt and pepper noise image, the neurons corresponding to noise pixels and its neighbor neurons corresponding to non-noise pixels will oscillate asynchronously. So the PCNN has good capability of noises detection. Accordingly, this paper presents a novel salt and pepper noise filtering method: first use a PCNN to detect noise pixels, then filter noise pixels by the extended median filtering (EMF) method proposed. Experimental results show that this method can filter out the salt and pepper noises in a noise image effectively. And comparing with MF, experimental results show distinct advantages also, especially for processing the noise images with high noise density.

The rest of this paper is organized as follows. Section 2 introduces the PCNN model briefly. In Sect. 3, a novel filtering method using PCNN is proposed. Section 4 presents the experimental results and discussions. Finally, we conclude in Sect. 5.

2 PCNN Model

Eckhorn model first described the behavior that the visual brain cortex neurons of cats burst synchronous pulse [2], and then Johnson [8] on the basis of the model presented PCNN model. Afterward, many researchers further improved PCNN model, and currently a popular PCNN model is described with the following Eqs. (1)–(5):

$$F_j(n) = S_j \tag{1}$$

$$L_j(n) = V_j^L \sum_k W_k^j Y_k(n-1) \tag{2}$$

$$U_j(n) = F_j(n)\big[1 + \beta_j L_j(n)\big] \tag{3}$$

$$\theta_j(n) = \theta_j(n-1)e^{-\alpha_j^T} + Y_j(n-1)V_j^T \tag{4}$$

$$Y_j(n) = step\,(U_j(n) - \theta_j(n)) = \begin{cases} 1, & U_j(n) > \theta_j(n) \\ 0, & otherwise \end{cases} \tag{5}$$

PCNN model consists of three parts: the receptive field, the modulation field and the pulse generator. In the receptive field, the neuron receives the input pulses from neighborhood neurons, and the pulses are transmitted through the feedback input channel F and link input channel L. W is the linking weight matrix of synapses between neuron and its neighbor neurons. S_j is the external stimulation signal (corresponding to the gray value in image processing), it acts as the output

of the channel F. The L channel signal attenuates exponentially, α^L is attenuation parameter and V^L is magnitude parameter. In the modulation field, $U_j(n)$ is the internal activity, which is obtained from the modulation between L channel and F channel, β is the link constant. In the pulse generator, when $U_j(n)$ is greater than the dynamic threshold $\theta_j(n)$, the corresponding neuron outputs a pulse, and then the threshold will increase to a high level. The $\theta_j(n)$ attenuate exponentially also, α^T and V^T is its attenuation parameter and magnitude parameter respectively.

3 Salt and Pepper Noise Filter Based on PCNN

3.1 Noise Detection Principal

In PCNN network, similar groups of neurons burst pulses synchronously. In addition, there are great differences in gray value between the noise pixels and non-noise pixels. Therefore, when an image is inputted as stimuli to a PCNN, the neuron corresponding to noise pixel and its neighbor neurons corresponding to non-noise pixels will oscillate asynchronously, namely the oscillation phases are different between the neurons with noise pixel and the neurons with non-noise pixel.

According to the synchronous or asynchronous oscillation between the neuron (i,j) and its neighbor neurons, we can judge whether the neuron is polluted by salt and pepper noise, this is to say the noise pixels can be located. In realization, threshold value N_0 can be set as 22 in the 5×5 widow. If the neuron (i,j) busts a pulse, but the total number of pulses from its neighbor neurons are less than N_0, the neuron (i,j) will be viewed a pixel polluted by noise, conversely it will be deemed as a pixel unpolluted. When the positions of noises have been located, a certain filtering algorithm such as MF method could be used to do specific filtering.

3.2 Filtering Method

For the noise pixel (i,j), the rational understanding is that the nearer non-noise pixels around it the more representative of its real gray value. And in order to protect the image edges effectively, a noise pixel can be estimated by the median of neighbor non-noise pixels. Consequently, a new filtering method is presented here. And the main idea is that the size of neighbor window for noise pixel (i,j) is expanded step by step, and the expansion is completed until the non-noises have been found in the window, and then replace the noise pixel (i,j) by the median of these non-noise pixels. This method is similar to the Median Filter (MF) except for not fixing the size of window. We may call it extended median filter (EMF).

As shown in Fig. 1, the largest window W_{max} of EMF is equal to 7. Firstly check whether there are non-noise pixels in window1 (the size is 3). If the non-noise pixels

Fig. 1 An example of
extending window in EMF
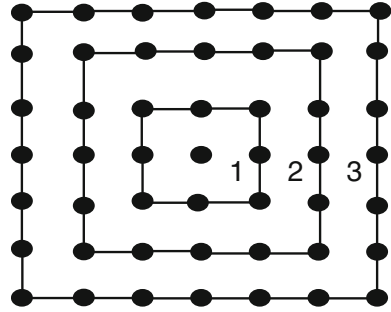
in the window exist, MF for the non-noise pixels will be carried on, and window expansion will be stopped. Otherwise, find the non-noise in window 2, etc.

The extended width should be restricted because it is unsuitable to estimate the pixel (i,j) by the non-noisy pixel too far from it. If the noisy pixel (i,j) cannot be filtered within the largest window W_{max} of EMF, it will be filtered by traditional MF with fixed the size of window.

3.3 Filtering Steps

According to the Eqs. (1)–(5), in a PCNN, the neurons with low gray value require more iterations to burst pulses, it leads to heavy calculation. And a more serious situation is that the neurons with gray 0 (they may be pepper noise) are unable to burst pulses, and these pepper noises can't be detected. Considering above problems, there are two steps to detect noises for a noise image. First detect the salt noises pixels corresponding to high gray: select the original image as the input to a PCNN, and then detect the pepper noises pixels corresponding to low gray: reverse the original image and select it as the input to the same PCNN. So the main steps of filtering the salt and pepper noises based on PCNN are as follows:
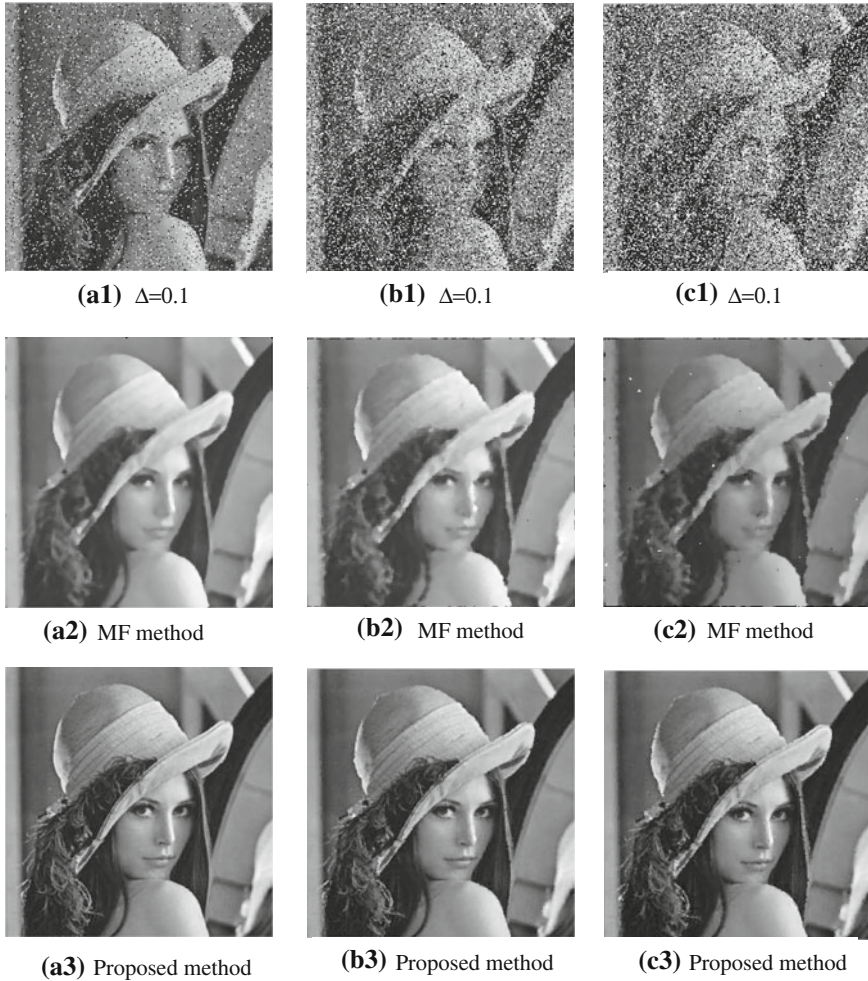
Step 1: Noise pixels detection. First, using a PCNN to detect the salt noise pixels with high gray, and then to detect the pepper noise pixels with low gray in the reversed image.

Step 2: Noises pixels filtering. On the basis of the noises detected in step 1, for a certain noise pixel, first using EMF to filter in largest window W_{max}. If successful, the filtering process will be completed. Otherwise, the noise pixel is filtered by MF with the 5×5 window.

4 Experiments and Results Discussion

To verify the effectiveness of the proposed filtering method, experiments studies are carried out on the 256×256 Lena gray image polluted by salt and pepper noise with different noise intensity Δ. In these simulations, the parameters of a
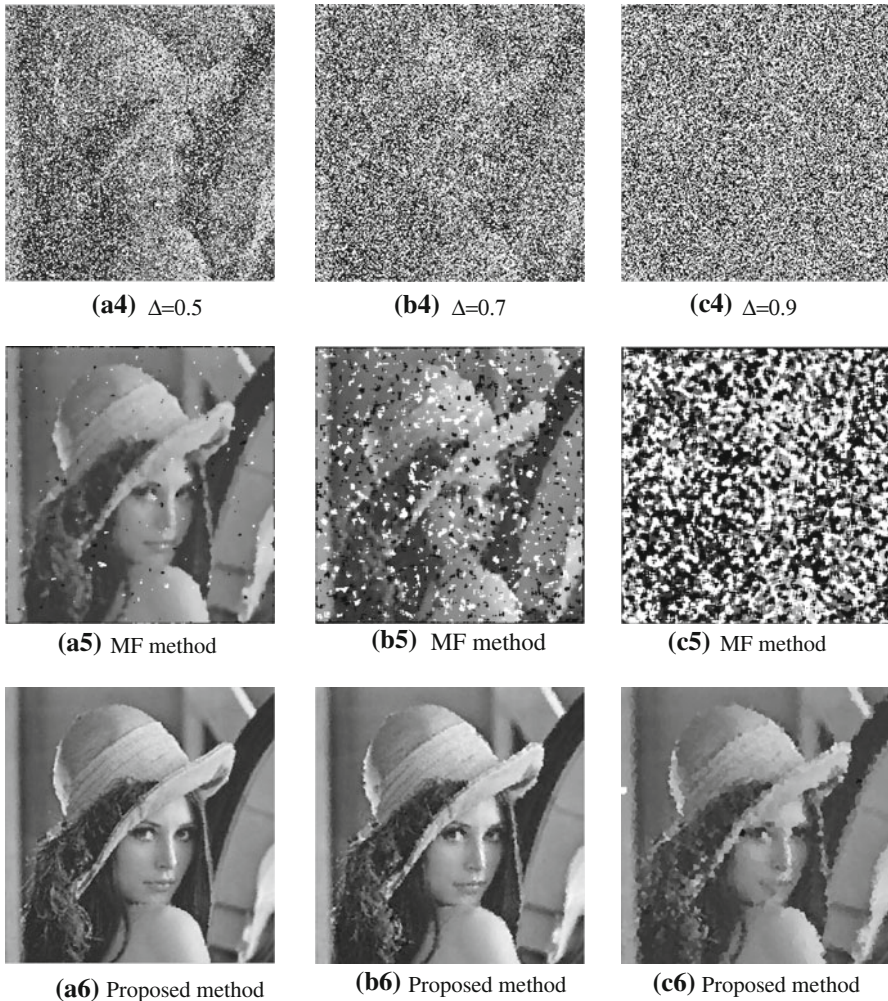
(a1) Δ=0.1   **(b1)** Δ=0.1   **(c1)** Δ=0.1

(a2) MF method   **(b2)** MF method   **(c2)** MF method

(a3) Proposed method   **(b3)** Proposed method   **(c3)** Proposed method

Fig. 2 The noise images and the images filtered. In a column (**a**, **b**, or **c**), the images are the noise image, the images filtered by MF and the image filtered proposed method. Noise intensity Δ ranges from 0.1 to 0.4

neuron are set as following: $W = [0.5\ 1\ 0.5;\ 1\ 0\ 1;\ 0.5\ 1\ 0.5]$, $\beta = 0.001$, $V^T = 10,000$, $\alpha^T = 0.0023$, $\theta(0) = 256$, $\alpha^L = 0.2$, $V^L = 1$. The maximum of iterations is 20 and the largest EMF window W_{max} is 11.

When noise intensity Δ is 0.1, 0.3 and 0.4, the filtering results by the proposed method and MF method are shown in Fig. 2. And the Fig. 3 shows the filtering results when noise intensity Δ is 0.5, 0.7 and 0.9.

It can be seen from Figs. 2 and 3 that the visual effect for the images filtered by the proposed method is better apparently than by MF method. For MF method, when the noise intensity Δ is greater than 0.4, the filtering effect is very bad. Some

(a4) Δ=0.5   **(b4)** Δ=0.7   **(c4)** Δ=0.9

(a5) MF method   **(b5)** MF method   **(c5)** MF method

(a6) Proposed method   **(b6)** Proposed method   **(c6)** Proposed method

Fig. 3 The noise images and the images filtered. In a column (**a**, **b** or **c**), the images are the noise image, the images filtered by MF and the image filtered proposed method. Noise intensity Δ ranges from 0.5 to 0.9

noise pixels which are not filtered correctly appear in the filtered images, and even we cannot get the main original image formation from filtered images; Even though when noise intensity Δ ranges from 0.1 to 0.4, there still exist very obvious fuzziness in the filtered images. But for the proposed method, when noise intensity Δ ranges from 0.1 to 0.4, the filtered images are very clear and there is hardly any visual difference compared with original un-polluted image; the filtering results become worse gradually with the increase of noise intensity, but even though Δ is

Table 1 PSNR of filtered images for different noise intensity Δ and filtering method

Noise Δ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
MF	28.56	27.24	25.68	24.02	21.92	17.85	13.73	10271	7.37
Proposed	40.11	36.48	34.05	32.41	30.55	28.94	27.67	26.03	23.83

Table 2 MSE of filtered images for different noise intensity Δ and filtering method

Noise Δ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
MF	90.5	122.9	175.9	257.7	418.1	1067	2759	6110	11565
Proposed	6.35	14.62	25.6	37.35	57.26	83	111	162	255

equal to 0.9, most of the information of the original un-polluted image still can be seen from the filtered images.

To verify the effectiveness of the proposed method objectively, we use objective criterions such as the peak signal to noise ratio (PSNR) and the mean square error (MSE) to evaluate the filtering quality of filtered images. Tables 1 and 2 show the PSNR and the MSE of the filtered images for different noise intensity and filtering method. Obviously, with the increasing of noise intensity, the PSNR by two methods both decrease. The PSNR by MF method reduces to 7.37 dB ($\Delta = 0.9$) from 28.56 dB ($\Delta = 0.1$), the reduction is about 21 dB. But the PSNR by the proposed method reduces to 23.83 dB from 40.11 dB, the reduction is about 16 dB. When Δ is equal to 0.1, the PSNR by the proposed method is about 20 dB higher than MF, and is approximately 17 dB higher even though $\Delta = 0.9$. Consequently, the filtering performance by the proposed method in terms of PSNR is far better than MF method with the noise intensity increasing.

As shown in table 2 that the MSE by these two methods also increases with the noise intensity increasing, so the filtered image quality deteriorates with the increasing of the noise intensity. The MSE by MF method increases to 11565 ($\Delta = 0.9$) from 90.5($\Delta = 0.1$), the increment is about 11474.5. But the MSE by the proposed method increases to 255 ($\Delta = 0.9$) from 6.35 ($\Delta = 0.1$), the increment is approximately 249. Apparently, the increment of MSE by MF method is much higher than the proposed method. In addition, when the noise intensity $\Delta = 0.1$, the MSE by the proposed method is about 84 lower than MF method, and is approximately 11310 lower when $\Delta = 0.9$. Thus, the proposed method has a greater advantage than MF method in terms of the MSE with the noise intensity increasing.

Generally, by subjective visual evaluation, the proposed method is significantly better than MF for noise images, especially for the images with larger noise intensity. Moreover in the aspects of objective evaluation criterions such as PSNR and MSE, the proposed method also has greater advantages than MF. Consequently, the salt and pepper noise filtering method proposed is reasonable and effective, and it is significantly better than MF method no matter from subjective evaluation or objective evaluation.

5 Conclusion

Based on the property of similar groups of neurons burst synchronous pulses in a PCNN, we realize the accurate detection of the noisy pixels in a noise image by a PCNN, and propose a novel filtering algorithm called extended median filter (EMF) that consider the similarity between noisy pixels and their neighbor non-noise pixels. And finally the paper proposed a novel filtering method for the salt and pepper noise image. Experimental results indicate that the proposed method has excellent filtering effectiveness for noise images with different noise intensity, and compared with MF method, the method has obvious and great advantages both in subjective visual evaluation and objective evaluation.

References

1. Guo M, Zhu M, Zhou XD (2011) A symmetrical orientation weighted mean filter for salt and pepper noise removing. Laser Infrared 41(11):1267–1272
2. Eckhorn R, Reitboeck HJ, Arndt M, Dicke P (1990) Feature linking via synchronization among distributed assemblies: simulation of result from cat visual cortex. Neutral Comput 2(3):293–307
3. Yang N, Chen HJ, Li YF et al (2012) Coupled parameter optimization of PCNN model and vehicle image segmentation. J Transp Syst Eng Inf Technol 12(1):3866–3875
4. Chai Y, Li HF, Guo MY (2011) Multifocus image fusion scheme based on features of multiscale products and PCNN in lifting stationary wavelet domain. Optics Commun 284(5):1146–1158
5. Liu F, Li J, Huang CY (2012) Image fusion algorithm based on simplified PCNN in nonsubsampled contourlet transform domain. Procedia Eng 29:1434–1438
6. Ji LP, Zhang Y (2008) A mixed noise image filtering method using weighted-linking PCNNs. Neurocomputing 71(13–15):2986–3000
7. Liu YM, Qin SY (2009) Novel adaptive denoising method for extreme noise based on PCNN. J Beijing Univ Aeronaut Astronaut 35(1):108–112
8. John JL, Ritter D (1993) Observation of periodic waves in a pulse coup led neural network. Opt Lett 18(15):1253–1255

Study of the Predictor Index System of Death Rate per Million-Ton Coal Based on Gray Relation

Huaping Zhou, Kelei Sun and Bojie Xiong

Abstract The thesis adopts gray correlation analysis methods, counting the country's coal mine safety data, taking the provinces (cities) as the basic unit, to identify the key impacting factors of death rate per million-ton coal by comparing and analyzing the samples of death rate per million-ton coal in different years. And it establishes the streamlined and reasonable predictor index system of death rate per million-ton coal.

Keywords Death rate per million-ton coal · Index system · Gray relation

1 Introduction

More than 90 % of China's coal mines are in underground. With the expanding of mining depth and mining scale, the difficulty of the mining and the safety risk of the coal mine will further aggravate. Coal mine safety risk is still severe from the safety accidents statistic data across the nation, Death rate per million-ton coal is an important measure of a region's coal mine safety production control assessment indicators in china. Many factors are related to death rate per million-ton coal. This paper has counted up the national safety production data, the provinces (municipalities) as the basic unit, by using the gray relation analysis method,

H. Zhou (✉) · K. Sun · B. Xiong
School of Computer Science and Engineering, Anhui University of Science and Technology, Huainan, Anhui, China
e-mail: hpzhou@aust.edu.cn

K. Sun
e-mail: klsun@aust.edu.cn

B. Xiong
e-mail: 873127611@qq.com

comparing the death rate per million-ton coal in different years, to find out the key factor that influences the death rate per million-ton coal, in order to set up a compact and reasonable death rate per million-ton coal prediction index system.

1.1 The Issued Way and the Decomposition Calculation Method of the Index at Present

In order to carry out the effective management of safety production, reduce casualties and the number of accidents, the State Council has developed "The decision of State Council about further strengthening safety production" in 2004 year. The production safety control indexes should be issued level by level and properly implemented in all the provinces and municipalities. Our country establishes the evaluation system of safe production control indexes [1]. The structure diagram of the control indexes issued level by level is shown in Fig. 1.

The decomposition calculation method is shown as follows:

1. About calculation base.

The number of deaths one year ago in various industries and fields will be as the calculation base for the control of assessment index. The number of deaths in coal mining enterprises is based on the reduction percentage in the previous year.

2. About calculation method

The calculation method is based on the ratio of the regional index value accounting for the country's index value correspondingly. The calculation formula is as follows:

$$Y_x = \frac{\frac{1}{5}\sum_{n=1}^{5} D_n^x}{\frac{1}{5}\sum_{n=1}^{5} G_n^x} \times A_x \qquad (1)$$
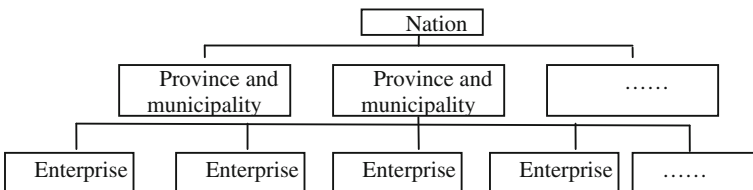
Fig. 1 Hierarchical diagram of control index

In formula (1):

Yx—The regional control index in that year; Dx—Each coal-producing administrative region's absolute index value in one year; Gx—The national absolute index value in one year; Ax—The national control index in that year; X—Various industries classification.

3. About adjustment factors

For certain areas' death rate per million-ton coal, if the actual value is too high, it also has to make some appropriate adjustments. Specifically, we make it compare to the national control assessment indexes, if it higher than 4 times of national death rate per million-ton coal, we set it as 4 times of national death rate per million-ton coal.

1.2 The Constitute of Impact Factors of Death Rate per Million-Ton Coal

When we forecast the provincial death rate of per million-ton coal, we must relate to the more specific and quantified indexes system. This paper refers to various mine safety regulations, such as "The Basic Conditions of The Coal Mine Production Safety Requirements", "Coal Mine Safety Regulations", "Safety Production Permit Implementation Measures", "Coal Mine Safety Evaluation Guidelines" and so on, considering the geological condition of our country's coal, mining technology, scientific and technological level, quality of employee and economical level's effect on coal mine safety, on this base, we have established this specific coal mine safety system [2]. It is shown in Fig. 2 many factors affect the death rate per million-ton coal, we are not able to list them, and they are not necessary for our prediction, we just select the key factors from them. According to the analysis of the entire coal mine safety system, 14 indicators meet this condition, they are as the primary indicators (Table 1).

2 Gray Relational Analysis

The general abstract systems, such as social systems, economic systems, agricultural systems and ecosystem, they include many kinds of factors, the result of a variety of factors determines the development trend of the system. Among these factors, we want to know which are major factors, which are secondary factors, which have more effects on the development of the system, which have fewer effects on the development of the system. All these problem are concerned in system analysis.
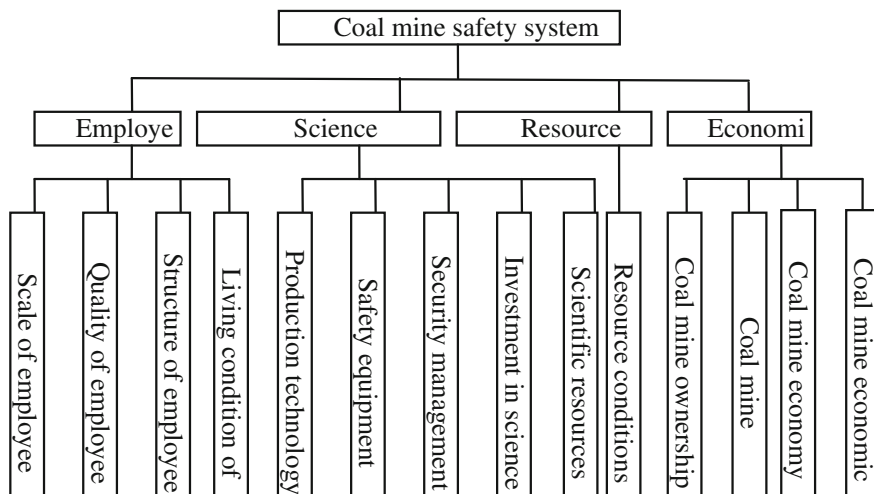
Fig. 2 Block diagram of coal mine safety system

Table 1 Primary indicator system of death rate of per million-ton coal

The influencing factor of death rate per million-ton coal	
X_1	The proportion of underground workers in employees
X_2	The proportion of technological works in employees
X_3	The proportion of rescue workers in underground workers
X_4	The rate of mining mechanization
X_5	The rate of fully mechanized mining
X_6	The rate of mechanized excavation
X_7	The proportion of stated-owned key coal production
X_8	The proportion of collectively owned coal production
X_9	The proportion of large-scale coal production
X_{10}	Full efficiency of raw coal
X_{11}	The average wage of all employees
X_{12}	Coal industrial output value accounted for the proportion of local secondary industry
X_{13}	The proportion of coal and gas outburst mines
X_{14}	The proportion of high-gas mines

The basic idea of gray relational analysis is based on the degree of similarity of the geometry sequence curve to determine whether they are close contact. If the curves are closer to each other, the degree of correlation between the corresponding sequences is greater; conversely, the degree is smaller [3]. Gray relational analysis is not only one of the important contents of the gray system theory, it is an important cornerstone for gray prediction, gray decision and gray control.

2.1 The Basic Characteristics of Gray Relational Analysis

1. Generality

In previous relational analysis, we always compare one factor to the other, gray relational analysis has broken down the limitation, it put all factors together, treating them as a system. A system contains a variety of factors, which have different importance and role to the system, but we can sort it according to the related degree, that is related sequence.

2. Non-uniqueness

The related degree is not unique, because a lot of factors should be considered when we calculate the related degree, such as the selection of the parent sequence, the size of the original data, the processing mode, and the value of the correlation coefficient.

3. Asymmetry

Various factors in the system are linked inextricably, it's hard to use a relationship to summarize, so the degree of association between two factors is not equal.

4. Dynamic

A system is changing over time, the relationship of the factors of the system will also change, and the related degree will change with it.

5. Order

One of the main targets of the analysis of gray correlation is time series. The change of data sequence and timing in the system will affect the original characteristic of sequence in the system.

2.2 The Model of Gray Relational Analysis

A reference number column X_0 with n elements can be represented as:

$$x_0 = [x_0(1), x_0(2), \ldots, x_0(n)] \tag{2}$$

Comparative sequence X_1, X_2,...,X_n can be represented as:

$$\begin{aligned}
x_1 &= [x_1(1), x_1(2), \ldots, x_1(n)] \\
x_2 &= [x_1(1), x_1(2), \ldots, x_1(n)] \\
&\quad \cdots \\
x_m &= [x_m(1), x_m(2), \ldots, x_m(n)]
\end{aligned} \tag{3}$$

The reference sequence and compare sequence can be seen as curves connected by data points. The essence of correlation analysis is to calculate the difference between them. The size of associated degree can be seen intuitively from the geometric shape. It is the area between reference curve and compare curve [4]. The process of calculation of associated degree is as fallows:

1. Pretreatment of raw data

The physical significant of each factor is different in the system, or the unit of measurement is different, which leads to the difference of dimensions of data, and sometimes, there are significance differences between the magnitudes of value. So it is inconvenient to make a comparison between different dimensions or different magnitudes, or even if we have done this, it is hard to get a correct result. In order to facilitate the analysis, at the same time, for the purpose of the equivalence and the same order of data, we need to make nondimensional data processing for the raw data before the comparison of each factor [5]. We often use mean value processing and initialization processing to make dimension normalized.

$X_i = (X_i(l), X_i(2), \ldots, X_i(n))$ is behavior sequence for X_i.

(a) Mean value processing

For a number column, we use its average value to be divided by all the number of it, the new number column we get is a mean value processed number column. We can assume that the original sequence is like this: $x_0 = (x_0(l), x_0(2), \ldots, x_0(n))$, the average of it is $\overline{x_0}$, y_0 is mean value processed x_0. It can be calculated like this:

$$y_0 = \{y_0(1), y_0(2), \cdots, y_0(n)\} = \left\{ \frac{x_0(1)}{\overline{x_0}}, \frac{x_0(2)}{\overline{x_0}}, \ldots, \frac{x_0(n)}{\overline{x_0}} \right\} \qquad (4)$$

(b) Initialization processing

For a number column, we use its first number to be divided by all the numbers, the new number column we get at last is an initialization processed number column.

The original sequence is like this: $x_0 = (x_0(l), x_0(2), \ldots, x_0(n))$, y_0 is initialization processed x_0. It can be calculated as follows:

$$y_0 = \{y_0(1), y_0(2), \ldots, y_0(n)\} = \left\{ \frac{x_0(1)}{x_0(1)}, \frac{x_0(2)}{x_0(1)}, \ldots, \frac{x_0(n)}{x_0(1)} \right\} \qquad (5)$$

2. The calculation of correlation coefficient

The degree of association between systems or factors is based on the similar degree of curves in geometry to determine whether they are close contact,

therefore, the difference between the curves can be used as a measure of the degree of association.

$$\xi_i(k) = \gamma(x_0(k), x_i(k)) = \frac{\min_i \min_k |x_0(k) - x_i(k)| + \xi. \max_i \max_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \xi. \max_i \max_k |x_0(k) - x_i(k)|} \quad (6)$$

x_i represents as compared curve, and x_0 represents as reference curve, they have different values in point K, they can be recorded as $x_i(k)$ and $x_0(k)$. The difference between them is $|x_i(k) - x_0(k)|$. It represents the correlation coefficient x_i with respect to of x_0 at time K. ξ is the distinguishing coefficient with a value between 0 and 1.

We calculate the minimum difference of the first level:

$$\Delta_i(\min) = \min_k |x_0(k) - x_i(k)| \quad (7)$$

$\Delta_i(\min)$ is an absolute difference for different value of K, the minimum value of $x_0(k) - x_i(k)$.

The minimum difference of the second level:

$$\Delta(\min) = \min_i(\min|x_0(k) - x_i(k)|). \quad (8)$$

It refers to the minimum one of $\Delta_1(\min), \Delta_2(\min), \ldots, \Delta_m(\min)$.

3. Calculation of associated degrees

The measure of the size of correlation between two systems or two factors can be called as associated degrees. The associated degree describes the relative changes in the factors during the development of the system, which is the relative of the change of size, direction and speed, etc. If the relative change of the two is consistent in the process of development, then we think that they have a high associated degree, otherwise, they have a low associated degree.

Formulation of associated degree between two curves (sequences):

$$\gamma_{0i} = \gamma(x_0, x_i) = \frac{1}{n} \cdot \sum_{i=1}^{n} \gamma[x_0(k), x_i(k)] \quad (9)$$

If we replace $\gamma[x_0(k), x_i(k)]$ to $\xi_i(k)$ and replace $\xi_i(k)$ to γ_i, then,

$$\gamma_i = \frac{1}{n} \cdot \sum_{i=1}^{n} \xi_i(k) \quad (10)$$

Table 2 Associated sequence of index system of death rate per million-ton coal

No.	2005 s.	d.	2006 s.	d.	2007 s.	d.	2008 s.	d.	2009 s.	d.	2010 s.	d.
1	x_8	0.5	x_2	0.51	x_8	0.44	x_{14}	0.47	x_8	0.55	x_8	0.52
2	x_{14}	0.49	x_8	0.49	x_{13}	0.44	x_8	0.46	x_2	0.52	x_{14}	0.5
3	x_2	0.46	x_{13}	0.47	x_{14}	0.43	x_{13}	0.43	x_{13}	0.51	x_{13}	0.5
4	x_{13}	0.46	x_{14}	0.46	x_4	0.42	x_2	0.42	x_{14}	0.48	x_2	0.48
5	x_{11}	0.45	x_4	0.42	x_6	0.41	x_6	0.41	x_{11}	0.45	x_6	0.46
6	x_4	0.41	x_6	0.41	x_2	0.41	x_5	0.4	x_6	0.42	x_4	0.43
7	x_6	0.39	x_5	0.37	x_{11}	0.38	x_{11}	0.38	x_4	0.38	x_{11}	0.41
8	x_5	0.37	x_{11}	0.37	x_5	0.37	x_4	0.35	x_5	0.33	x_5	0.39
9	x_{10}	0.36	x_{10}	0.35	x_1	0.36	x_{10}	0.32	x_1	0.31	x_{10}	0.36
10	x_3	0.24	x_3	0.29	x_3	0.21	x_7	0.25	x_3	0.22	x_3	0.24
11	x_1	0.22	x_9	0.27	x_{12}	0.2	x_9	0.23	x_{10}	0.21	x_1	0.21
12	x_9	0.21	x_7	0.26	x_7	0.18	x_3	0.22	x_7	0.19	x_9	0.18
13	x_{12}	0.21	x_{12}	0.24	x_{10}	0.18	x_{12}	0.22	x_{12}	0.17	x_{12}	0.15
14	x_7	0.18	$x1$	0.22	x_9	0.16	x_1	0.19	x_9	0.14	x_7	0.14

s. sequence; *d.* degree

3 Associated Analysis of Death Rate per Million-Ton Coal

According to the above calculation method, we sort the index system of death rate per million-ton coal since 2005 in accordance with gray relational analysis, the result is shown in Table 2 [6, 7].

(1) We can see the associated analysis from 2005 to 2010, the associated sequences of the 14 factors which affect death rate per million-ton coal have a little difference, but in general, they are in accordance with each other. (2) Although the associated sorting of the first 9 factors changes individually, as a whole, they are in consistent, the influencing factors of the first 9 factors are the same. (3) The first 10 influencing factors are no longer the same, and associated degree of the last 5 influencing factors is small.

4 Conclusion

Firstly, the paper describes the composition of China's control assessment indexes in coal mines, the calculation method of control indexes of death rate per million-ton coal in China and the method of indexes issued and decomposed. Secondly, it discusses the influencing factors of death rate per million-ton coal on the basis of creating a regional coal, and then it establishes the mode of gray correlation analysis. With this method, it selects 9 factors which have a high associated degree with death rate per million-ton coal and reflect the death rate per million-ton coal.

References

1. Zhang X, Zhang R (2009) Analysis theory research on regional coal mine safety system. China Coal 10:75–79
2. Li Y, Zhang X (2009) Construction of index system for regional safety production system. China coal industry 33:243–247
3. Jiang L, Qi YB (2011) The evaluation on ecological agricultural development in Sichuan province on the basis of grey relational analysis. Asian Agric Res 5:188–196
4. Zhang XF, Qi YB (2011) The grey correlation analysis on the relationship between the rural finance and the rural economical growth in Sichuan province. Asian Agric Res 1:323–331
5. Ye XL, Yao ZZ (2012) Study on evaluation of the rainstorm disaster in Fujian province based on spectral clustering model with grey correlation analysis. Meteorol Environ Res 6:401–409
6. Coal Industry Annual Statistics Report. The National Security Supervision Administration 2000–2011
7. The national coal industrial safety production accident statistic and analysis report. Statistics Division of the State Administration of Coal Mine Safety Supervision Bureau 2000–2011

Multi-Objective Optimization of Post-Disaster Emergency Resources Scheduling Using Multiple Ant Colony Systems Optimization Algorithm

Renqiang Wen, Shaobo Zhong and Bin Zhang

Abstract The multi-objective optimization mathematical model for post-disaster emergency resources scheduling was built. This model took into account multiple demand centers, multiple supply centers, multiple kinds of resources, and supply centers cooperating with each other in providing resources. The reliability of scheduling routes was taken into account to enhance the practicability. A new approach using multiple ant colony systems optimization algorithm was proposed to solve the model. The elitist strategy was imported into the global pheromone update strategy to guide exchanging and sharing information among multiple ant colony systems, and it improved the efficiency in searching global no-inferior solutions. It provided a practical approach to integrate resources location-allocation problem and scheduling routes planning problem together. The case was presented to verify the validity of the model and algorithm.

Keywords Multiple ant colony systems · Ant colony optimization · Optimization resource scheduling · Multi-objective optimization

1 Introduction

Resource scheduling is a Location Routing Problem (LRP) [1] which integrates Location-Allocation Problem (LAP) and Vehicle Routing Problem (VRP). Post-disaster emergency resource scheduling shows apparent weak-economy character.

R. Wen (✉)
National Marine Hazard Mitigation Service, State Oceanic Administration, Beijing, China
e-mail: wenrenqiang@hotmail.com

S. Zhong · B. Zhang
Department of Engineering Physics, Tsinghua University, Beijing, China
e-mail: zhongshaobo@tsinghua.edu.cn

The major difference between post-disaster emergency resource scheduling and commercial logistics is that efficiency and safety are both important and are taken into account together in the model. LRP is a NP-hard problem. Many LRP models have been developed to solve practical problems, such as parcel delivery [2, 3], medical evacuation [4], bill delivery [5], optical network design [6], and telecom network design [7].

Related research work in the field of emergency resource scheduling started late. But now, more and more researchers focus on this field. Yuan and Dingwei [8] proposed a two-objective model with total travel time minimization object and path complexity minimization object. The chaos, panic and congestions were taken into account in the parameter of time. An ant colony optimization algorithm was proposed. Susheng and Yan [9] proposed a model to solve the optimized resources allocation problem in large-scale public emergency. The optimized object is that advance the earliest start time in emergency response, and a heuristic algorithm was designed. Yi and özdamar [10] proposed an integrated location-distribution model which involves dispatching commodities to distribution centers, evacuation, and transfer of wounded people. It was a mixed integer multi-commodity network flow model, and a two-stage procedure was proposed. Ozdamar and Ekinci [11] proposed a logistics planning model which addressed the dynamic time-dependant transportation problem that needed to be solved repetitively at given time intervals during ongoing aid delivery. It was a hybrid problem that integrated the multi-commodity network flow problem and the vehicle routing problem. In the solution approach, the sub-problems were coupled with relaxed arc capacity constrains using Lagrangean relaxation. Fiedrich and Gehbauer [12] proposed a dynamic optimization model to find the best assignment of available resource to operational areas. The model used detailed descriptions of the operational areas and of the available resources to calculate the resource performance and efficiency for different tasks related to the response. At last, heuristics Simulated Annealing (SA) and Tabu Search (TS) were used to solve the model.

Literature reviewing indicates that research works on the integrated problem which includes resource location-allocation and scheduling path selecting with multiple demand centers, multiple supply centers, multiple kinds of resources, and supply centers cooperating with each other in providing resources are insufficient. Therefore a new model and algorithm have to be developed.

2 Model

2.1 Parameters

A: resource demand centers; B: resource supply centers; C: transfer sites; Γ_{ij}: the shortest travel time between node N_i and node N_j; A_i: resource demand center i; B_j: resource supply center j; C_k: transfer site k; P: amount of resource type; Q: resource requirement sets; Q_{ip}: the quantity of resource requirement of type p from

A_i; D: resource supply sets; D_{jp}: the quantity of resource supply of type p from B_j; S_{ij}: one route connects N_i and N_j; S_i: an arc of S_{ij}; $\Gamma(S_{ij})$: travel time on S_{ij}; $R(S_{ij})$: reliability of S_{ij}; $R(S_i)$: reliability of S_i; $E(C_k)$: delay on C_k; ΔQ_{ij}: the quantity of unsatisfied resource of type j on A_i; x_{ji}^p: the quantity of resource of type p from B_j to A_i. $S(A_i)$: the arcs of resource dispatching route from resource supply centers to A_i.

2.2 Mathematical Formulation

$$\max V_1 = \max \sum_{i \in I} R(S(A_i)) \tag{1}$$

$$\max \min V_2 = \max \min_{i \in I} R(S(A_i)) \tag{2}$$

$$\min V_3 = \min \sum_{i \in I} \left(\sum_{S_i \in S(A_i)} \Gamma(S_i) + \sum_{C_k \in S(A_i)} E(C_k) \right) \tag{3}$$

$$\min \max V_4 = \min \max_{i \in I} \left(\sum_{S_i \in S(A_i)} \Gamma(S_i) + \sum_{C_k \in S(A_i)} E(C_k) \right) \tag{4}$$

$$\min \max V_5 = \min \max_{i \in I} \sum_{r=1}^{P} \Delta Q_{ir} \tag{5}$$

$$\sum_{i=1}^{n} x_{ji}^r \le D_{jr}, \forall j, 1 \le r \le P \tag{6}$$

(1) Object V_1 maximizes the sum of all dispatching routes' reliability. Object V_2 maximizes the minimal reliability of the dispatching routes. (2) Object V_3 minimizes the total time cost of all dispatching routes. Time cost includes travel cost and transfer delay. It can advance the time of dispatching mission completion. Object V_4 makes resource demand centers receive resources as early as possible. (3) Object V_5 minimizes the quantity of unsatisfied resources. (4) Formulation (6) restricts the quantity of each type of resources from resource supply centers cannot exceed its reserve.

3 Algorithm Design

The two sub-problems in the model affect each other. So, we design special ant system for each sub-problem to search their local optimized solutions. The heuristic information, pheromone information and pheromone update strategy are different

in each ant system. According to the relationship between the two sub-problems, design the global pheromone update strategy under the guide of Pareto non-inferior solution principle [13]. Global pheromone update strategy can guide the two ant systems to search the global optimized no-inferior solution together.

3.1 Multiple Ant Systems

(1) **Emergency Resource Location-Allocation Ant System (LA System)**

Heuristic information $\eta_{ij}(\eta_{ij} = \eta_{ij}^1\eta_{ij}^2)$ includes two parts defined as follows:

$$\eta_{ij}^1 = \frac{1}{\Gamma_{ij} + \Delta\Gamma} \tag{7}$$

$$\eta_{ij}^2 = \frac{D_j'}{Q_i'} = \frac{\sum\limits_{r=1}^{P}\left(\theta_r D_{jr}'\right)}{\sum\limits_{r=1}^{P} Q_{ir}'} \tag{8}$$

where $\Delta\Gamma$ is the delay of changing transport mode, Q_{ir}' is the quantity of unsatisfied resources of type r of A_i, D_{jr}' is the quantity of remains of resources of type r of B_j is a binary variable (If $Q_{ir}' > 0, \theta_r = 1$, otherwise $\theta_r = 0$). Local pheromone update rule is stated as follows:

$$\tau_{ij}(t+1) = (1 - \xi)\tau_{ij}(t) + \xi\tau_0, \ 0 < \xi < 1 \tag{9}$$

where t is iteration step number, ξ is evaporation rate, τ_0 is the initial value.

(2) **Emergency Resource Dispatching Path Selection Ant System (DPS System)**

Heuristic information $\eta_{ij}(\eta_{ij} = \eta_{ij}^1\eta_{ij}^2)$ includes two parts defined as follows:

$$\eta_{ij}^1 = \frac{1}{\Gamma_{ij} + \theta(C_j)E(C_j) + \theta(C_i)E(C_i)} \tag{10}$$

$$\eta_{ij}^2 = R(S_{ij}) \tag{11}$$

where $\theta(C_j)$ is a binary variable (If C_j is transfer site, $\theta(C_j) = 1$, otherwise $\theta(C_j) = 0$). Local pheromone information update is performed on each object. The update rule is the same as stated in formulation (9).

(3) **Elite Document Update**

χ^* is stated as global optimized solution constructed by Moy^* (local optimized solution from LA System) and χ (local optimized solution from DPS System). Once iteration ends, remove the solution dominated [13] by current global optimized solution out from elite document, and perform pheromone update as follows:

$$\tau_{ij}^k = \tau_{ij}^k - t\delta \qquad (12)$$

where δ is a small positive number. If $\tau_{ij}^k < \tau_0$, set $\tau_{ij}^k = \tau_0$.

(4) **Global Pheromone Update Strategy**

The global pheromone update strategy is stated as follows:

$$\tau_{ij}^k(t+1) = (1 - \xi)\tau_{ij}^k(t) + \xi\Delta\tau_{ij}^k \qquad (13)$$

$$\Delta\tau_{ij}^k = \begin{cases} 10, (i,j) \in Best_1^k \\ 5, (i,j) \in Best_2^k \\ 0, \text{otherwise} \end{cases} \qquad (14)$$

where $\Delta\tau_{ij}^k$ is pheromone accumulation. The best optimized solution on object k is stated as $Best_1^k$, and the second-best optimized solution is stated as $Best_2^k$.

3.2 Algorithm to Solve the Model

The nodes visited by ant k are stored in $Tabu_k$. Local solutions searched by ant k are stored in Moy_k and χ_k. The algorithm is described as follows:

Step 1: Initialize parameters. $Iter = 0$.
Step 2: If iteration exceeds the maximal step number, exit algorithm.
Step 3: Generate LA System: initialization, M ants, $k = 0$. Select ant k.
Step 4: Initialize $Tabu_k$ and Moy_k.
Step 5: Select A_i, construct local solution, and update $Tabu_k$ and Moy_k.
Step 6: If all the elements in A have been visited, go to STEP 7. Otherwise, $i = i + 1$, go to STEP 5.
Step 7: Update local pheromone. $k = k + 1$. If $k < M$, go to STEP 4, otherwise exit. Get local optimized solution Moy^*.
Step 8: Generate DSP System: Initialization, M ants, $k = 0$. Select ant k.
Step 9: Initialize $Tabu_k$ and χ_k. Select $Moy_k^*(Moy_k^* \in Moy^*)$.
Step 10: Select $Moy_k^{A_i}(Moy_k^{A_i} \in Moy_k^*)$.
Step 11: Construct resource dispatching route of A_i. Update χ_k.
Step 12: If all demand centers included in Moy_k^* have been visited then go to STEP 13. Otherwise, $i = i + 1$, go to STEP 10.

Step 13: Add χ_k to χ, and update local pheromone. $k = k + 1$. If $k < M$ then go to STEP 9, otherwise go to STEP 14.

Step 14: Construct the global optimized solution χ^*. Perform elite document update and global pheromone information update.

Step 15: $Iter = Iter + 1$, go to STEP 2.

4 Case Study

In Fig. 1, demand centers: $I = \{I1 \sim I32\}$, resource depots: $C = \{C1 \sim C5\}$, transfer sites: $K = \{K1 \sim K11\}$, supply centers: $J = \{J1 \sim J25\}$, and road nodes: $N = \{N1 \sim N27\}$. Resource requirement and reserve are listed in Table 1.

Reliability of routes: $J9 \rightarrow K8$ is 0; $J25 \rightarrow I31$ is 0.2; $J19 \rightarrow K5$ is 0.1; $J21 \rightarrow I27$ is 0; $J20 \rightarrow I17$ is 0; $J11 \rightarrow I15$ is 0.1; $J3 \rightarrow I13$ is 0.1. Parameters initialization: $M = 20, Iter_{max} = 100, q_0 = 0.9 \sim 0.4, \xi = 0.1, \delta = 0.01, \tau_0 = 1, Delay = 1$ hour. LA System: $\alpha = 1, \beta = 1$. DSP System: $\alpha = 1, \beta = 3$. One of the global non-inferior optimized solutions is listed in Table 2 and Fig. 1.

Case shows that, when resources are insufficient on $J14$ and $J20$, algorithm can dispatches resources from $J13$ and $C5$ respectively. (1) Two routes from $J13$ to $I3$: A is $J13 \rightarrow K9 \rightarrow K8 \rightarrow I3$, B is $J13 \rightarrow N24 \rightarrow K9 \rightarrow K8 \rightarrow I3$. Although route A is shorter than route B, B is superior on the time object than route A. (2) Two routes from $J11$ to $I15$: C is $J11 \rightarrow J4 \rightarrow I15$, D is $J11 \rightarrow I15$. Reliability of C is 1, and D is 0.1. Although route C is longer than D, C is superior on the reliability object.
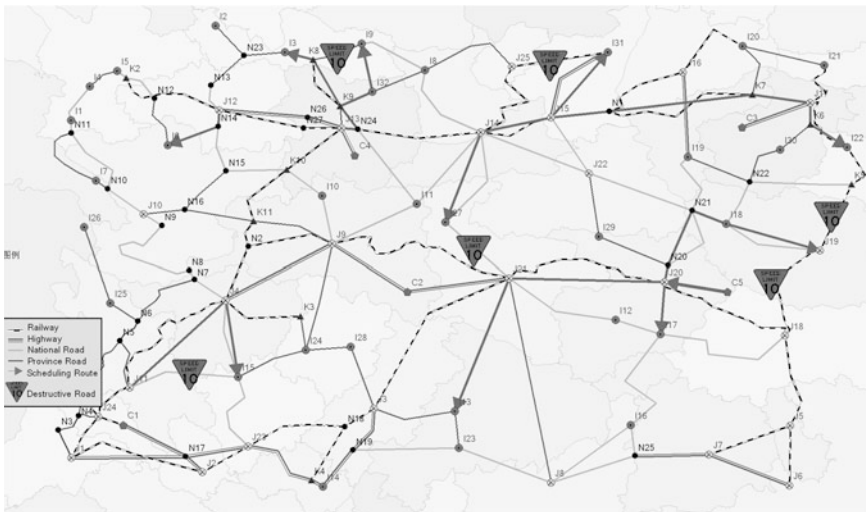

Fig. 1 Resource scheduling routes are visualized by GIS technology

Table 1 The quantity of resource requirements in one week and resource reserve

Demand center	Supply center	Number of people	Food (Kg)	Tent	Water (L)
$I3$	–	6,000	42,000	750	252,000
$I6$	–	7,100	49,700	888	298,200
$I9$	–	7,300	51,100	913	306,600
$I13$	–	7,900	55,300	988	331,800
$I15$	–	5,900	41,300	738	247,800
$I17$	–	7,600	53,200	950	319,200
$I19$	–	9,100	63,700	1,138	382,200
$I22$	–	11,300	79,100	1,413	474,600
$I27$	–	9,180	64,260	1,148	385,560
$I29$	–	7,300	51,100	913	306,600
$I31$	–	9,892	69,244	1,237	415,464
–	$J11$	–	150,360	2,685	902,160
–	$J12$	–	128,800	2,300	772,800
–	$J13$	–	133,700	2,388	802,200
–	$J14$	–	207,004	3,697	1,242,024
–	$J20$	–	208,600	3,725	1,251,600

Table 2 Resources location-allocation result and dispatching routes

From	To	Food (Kg)	Tent	Water (L)	Dispatching route
$J11$	$I15$	41,300	738	247,800	$J11 \rightarrow J4 \rightarrow I15$
$J12$	$I6$	49,700	888	298,200	$J12 \rightarrow N14 \rightarrow I6$
$J13$	$I3$	42,000	750	252,000	$J13 \rightarrow N24 \rightarrow K9 \rightarrow K8 \rightarrow I3$
$J13$	$I9$	51,100	913	306,600	$J13 \rightarrow K9 \rightarrow I32 \rightarrow I9$
$J14$	$I22$	79,100	1,413	474,600	$J14 \rightarrow J15 \rightarrow N1 \rightarrow K7 \rightarrow J17 \rightarrow K6 \rightarrow I22$
$J14$	$I27$	64,260	1,148	385,560	$J14 \rightarrow I27$
$J14$	$I31$	69,244	1,237	415,464	$J14 \rightarrow J15 \rightarrow I31$
$J13$	$I14$	5,600	100	33,600	$J13 \rightarrow N24 \rightarrow J14$
$J20$	$I13$	55,300	988	331,800	$J20 \rightarrow J21 \rightarrow I13$
$J20$	$I17$	53,200	950	319,200	$J20 \rightarrow J18 \rightarrow I17$
$J20$	$I29$	51,100	913	306,600	$J20 \rightarrow N20 \rightarrow I29$
$J20$	$I19$	63,700	1,138	382,200	$J20 \rightarrow N20 \rightarrow N21 \rightarrow I18 \rightarrow I19$
$C5$	$J20$	14,700	263	88,200	$C5 \rightarrow J20$

5 Conclusion

It built multi-objective optimization mathematical model for post-disaster emergency resources scheduling. This model took into account multiple demand centers, multiple supply centers, multiple kinds of resources, and supply centers cooperating with each other in providing resources. It proposed a new approach using multiple ant colony systems optimization algorithm to solve the model. The elitist strategy was imported into the global pheromone update strategy to guide exchanging and sharing information among multiple ant colony systems, and

improved the efficiency in searching global no-inferior solutions. It provided a practical approach to integrate resources location-allocation problem and scheduling routes planning problem together. The case showed that the algorithm can deals with large-scale, complex network well.

References

1. Nagy G, Salhi S (2007) Location-routing: issues, models and methods. Eur J Oper Res 177(2):649–672
2. Bruns A, Klose A, Stahly P (2000) Restructuring of Swiss parcel delivery services. OR Spektrum 22(2):285–302
3. Wasner M, Zäpfel G (2004) An integrated multi-depot hub-location vehicle routing model for network planning of parcel service. Int J Prod Econ 90(3):403–419
4. Chan Y, Carter WB, Burnes MD (2001) A multiple-depot, multiple-vehicle, location-routing problem with stochastically processed demands. Comput Oper Res 28(8):803–826
5. Lin CKY, Chow CK, Chen A (2002) A Location-routing-loading Problem for Bill Delivery Services. Comput Ind Eng 43(1–2):5–25
6. Lee Y, Kim S-I, Lee S, Kang K (2003) A location-routing problem in designing optical internet access with WDM systems. Photon Netw Commun 6(2):151–160
7. Billionnet A, Elloumi S, Grouz-Djerbi L (2005) Designing radio-mobile access networks based on synchronous digital hierarchy rings. Comput Oper Res 32(2):379–394
8. Yuan Y, Dingwei W (2009) Path selection model and algorithm for emergency logistics management. Comput Ind Eng 56(3):1081–1094
9. Susheng W, Yan W, Jian S (2007) An optimized emergency resources allocation algorithm for large-scale public emergency. In: Proceedings of the Sixth International Conference on Machine Learning and Cybernetics (ICMLC 2007), pp. 119–123. IEEE Computer Society, Piscataway (2007)
10. Yi W, Özdamar L (2007) A dynamic logistics coordination model for evacuation and support in disaster response activities. Eur J Oper Res 179(3):1177–1193
11. Ozdamar L, Ekinci E, Kucukyazici B (2004) Emergency logistics planning in natural disasters. Ann Oper Res 129(1–4):217–245
12. Fiedrich F, Gehbauer F, Rickers U (2000) Optimized resource allocation for emergency response after earthquake disasters. Saf Sci 35(1–3):41–57
13. Karl D, Walter J, Richard F, Christine S, Christian S (2004) Pareto ant colony optimization: a metaheuristic approach to multiobjective portfolio selection. Ann Oper Res 131(1–4):79–99

Research on Modeling and Simulation of Virtual Miner Agent in Coalmine Virtual Environment

Jingjing Tang, Zhixiang Yin, Liguo Qu and Chaoli Tang

Abstract For the underground coal mines safety behavior analysis, the behavior modeling and motion control technology of virtual miner agent is studied under the virtual environment. Based on the theory modeling method of agent, the system structure and its formal description of virtual miner with decision ability is built and is used in the virtual environment, which is integrated with the perception, information processing, learning, behavior, planning, decision-making, and knowledge base and so on, and it can effectively realize the behavior control of virtual miner, as well as typical operation of the behavior simulation of underground, and it can generate lifelike personification behavior. The control of the complex behavior of the virtual miners' agent and interactive simulation of virtual miner in the coalmine virtual environment are both realized in the PC using the object-oriented technology in this paper. So we build a behavior with realistic of underground of virtual miner in order to provide technical reference

Keywords Virtual simulation · Virtual miner agent · Coalmine virtual environment (CMVE) · Behavioral modeling

1 Introduction

Intelligent virtual human technology has been widely used in games, military training, network virtual meetings, distance education and other fields, and different applications focus on different attribute requirements of virtual human, such as autonomous behavior, interactive behavior, personalization emotional and

J. Tang · Z. Yin · L. Qu · C. Tang (✉)
Anhui University of Science and Technology, Huainan 232001 Anhui, China
e-mail: chaolitang@163.com

J. Tang
e-mail: 595455103@qq.com

cognitive planning, environmental adaptability and so on, how to build effective modeling of virtual people and simulate realistic of human behavior due to the complexity of human behavior itself, it is stilling one of the most challenging problem in current research of virtual human [1].

It is very extensive that studied the modeling of virtual human under virtual environment [2], which have no unified standard, and different areas of research core are not the same. In the coal areas, it has been some specific applications from the technology of virtual environment that has used in design of mine virtual [3], an operations simulation of virtual mine [4, 5], the training of virtual safety [6], safety assessment of underground [7], disaster accident simulation of mine [8] and so on, and which increasingly shows prospects of broad application [9, 10], However, how to explain the complex behavior of the underground virtual environment, especially the various acts of the underground virtual miner, which is the key to building a really and effectually the virtual environment of underground. This paper will combine the environment of underground and the technology of virtual human to research the modeling of behavior of virtual miner agent in virtual environment and the technology of motion control. The virtual environment intelligent of virtual miner modeling is embodied, virtual miner model is constructed with perception, movement, behavior, cognitive ability which can make the miners simulation more intelligent, and realize behavioral controlling of virtual miner effectively, as well as underground typical job of the behavior simulation; study virtual miners' own behavior and interaction behavior to build underground virtual environment with the realistic of behavior, and effective method is offord to the study of the underground production complex system.

2 Agent Basic Concepts

Agent evolved from the field of human intelligence, different schools for the agent have different definitions, of the most widely recognized to agent under "weak definition" and "strong definition" of Wooldridge et al. [11], "weak definition" refers agent with independent ability, social skills, the ability to respond to hardware and software and the ability to pre-action system; "strong definition" of gent not only has all the characteristics, but also has the characteristics of human beings are capable of knowledge, belief, purpose, and obligations and so on.

In general, agent refers to those activities in a complex and dynamic environment, self-aware environment, and be flexible enough to take action and to achieve a series of pre-set goals or tasks computing entity or program. The agent generally has autonomy, reactivity, initiative, adaptability, and coordination, social, learning and other properties.

Agent's structure refers to the relationship of the information and control system between each agent, as well as their distribution morphology. The agent of structure is the main basis for the design agent. Agent structure has mainly three types: the structure of Reactive, the structure of cognitive, the structure of hybrid [11, 12].

3 Design and Modeling of the Virtual Miner Agent

3.1 Functionality Requirements of the Virtual Miner Agent

To coalmine virtual environment, the virtual miner agent can become the incarnation of the participants receive instruction execution emulation action, the performance of the physical skills needed by the simulation; participants simulation tasks decision-makers, navigation in a virtual mine assessment situation and make decisions, issue instructions to control the behavior of virtual miner agent. At the same time, the virtual miner agent should be able to self-perception of the coalmine virtual operating environment, and in accordance with current goals, planning, and implementation of various simulation tasks. Whole coalmine virtual environment, the virtual miner agent according to the changes in the perception of the environment and its own internal attributes of external events to make real-time, reasonable reaction, such as showing fear of dangerous source, to reflect their personality traits, increased immersive virtual environment. Therefore, in order to realistically simulate the behavior of underground miners, we need to construct intelligent virtual must have the overall characteristics of the people, should be able to perceive the virtual environment, analysis of the environmental situation, decision-making and freedom of movement, while thinking, emotional and mental state and behavior ability as an incarnation of the participants in the underground virtual environment, instead of participants perform simulation tasks, reflecting the participants' decision-making ability.

3.2 Formal Definition of the Virtual Miner Agent

Definition: *Agent*. Agent is the one consisting of a five-tuple: the identity of the agent, agent attribute set, the inference engine inference, knowledge base, communicate.

Agent: : = < AgentID, Attribs, Inference, Knowledge, Communicate >,

which, AgentID is the only identity mark for agent in a virtual environment; Attribs is the collection of properties for the Agent; Inference is the reasoning of engine, and it is a core part of the agent. When the agent trigger condition is met, call the corresponding behavior of the function, operation to activate the appropriate behavior. Reasoning to specify the name of reasoning called behavior rules, behavior trigger conditions, and the behavior of time constraints. Knowledge defined in the dynamic target constraints, rules, decision-making knowledge and new concepts. Communicate is a collection of all the messages, the message is the medium of communication between the agent in the virtual environment, and the communication and collaboration.

3.3 Structure Model of the Virtual Miner Agent

Virtual miner agent fuzzy decision rules as a decision-making mechanism for the realization fuzzy judge and fuzzy decisions, virtual miner agent must have a perception module for data acquisition and information processing capabilities with the knowledge base of the fuzzy decision-making function of the decision-making module to store information for the decision rules base. The Reactive Agent Cognitive or agent apparently unable to realize the functional requirements of the virtual miner agent, therefore, based on the hybrid agent structure to build virtual miner agent architecture, structure shown in Fig. 1, as shown in Fig. 1 full advantage of the hybrid agent structure the perception and behavior of the ability to achieve the miners agent, and the use of the hybrid agent communication module implements the decision-making information.

In Fig. 1, the virtual miner agent perception real time the external virtual environment information; motion controller is used to update the posture and position of the virtual miner interaction with the virtual environment; coordination controller is responsible for the coordination of the entire agent running, to achieve complex behavior control; learning systems and knowledge systems with the outstanding performance of the intelligent agent, and the complete agent self-learning and complex reasoning and decision-making; role of the reactor in order to make the agent to respond quickly to emergency or simply; stored in the internal properties of the parameters associated with the virtual miner personality characteristics, including physiological attribute and psychological variables, such as
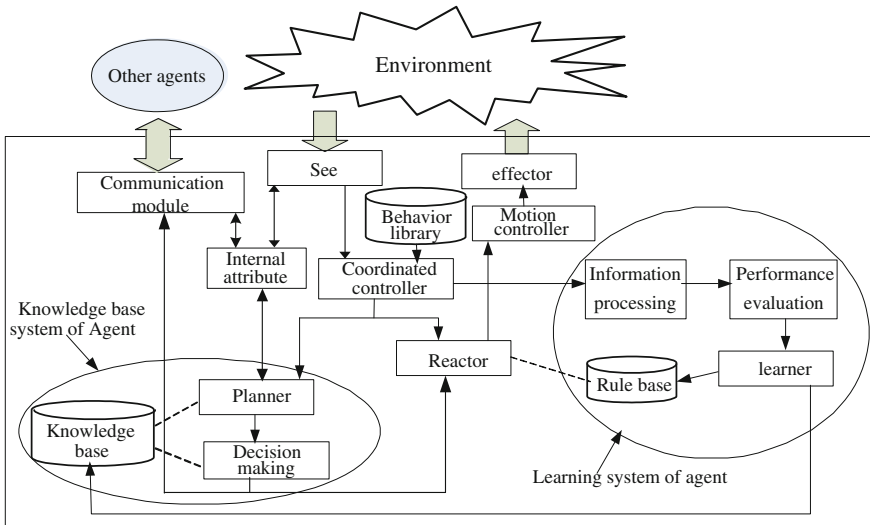

Fig. 1 Structure of virtual miner agent

fatigue, fear, neglect; planner based on the static model of the virtual environment of the knowledge base, dynamic sensing information field knowledge and task planning and planning results after the motion controller.

4 Virtual Miner Agent Implementations

Virtual miner agent architecture requirements, according to Fig. 1, the constructed virtual miner's agent model shown in Fig. 2. This paper discussed in detail below the virtual miner's agent model functions modules and work process.

4.1 Perceptron and Effector

Perceptron and effector is the interface of virtual miner agent and external environment, which has certain information processing function. Virtual miner agent can obtain current state information through using perceptron, and can abstract the state of other agent and environment. Perceptron sends the abstracted information to coordination controller. Then effector gives out corresponding action according to the input motion instruction to influence the external environment.

4.2 Coordinated Controller

Coordination controller is responsible for the virtual miner agent coordination run, Perception found that the status of the coal mine production environment changes (such as underground gas concentration in excess of Roadways ineffective) or

Fig. 2 Virtual miners agent model

received other agent's task request, coordinated controller based on the information types and coordinated control rules interpreted and classified information, and assigned to the unit of work.

4.3 Learning Device

Learning device is a virtual miner agent intelligent performance. When the coordination controller detects the information is new knowledge or mine production environment feedback information, virtual miner agent through information processing and performance evaluation module, information handling, processing and analysis, the formation of knowledge to join the Knowledge Base, or the rules of the library to be updated. The learning enables miner agent storage experience with the simulation, the same or even completely different the next simulation environment, they use these experiences to change their behavior, the results of simulation also change in order to achieve the miner in the coal mine who is in the environment of virtual agent simulation of adaptive behavior.

4.4 Decider Unit

Virtual miner agent decision-making process with uncertain reasoning methods and case-based reasoning technology, the miners' behavior to predict and guide its next step behavior, the randomness of the virtual miner agent behavior information, fuzzy, time-varying, incomplete and inconsistencies, and so, in the decision-making process in the organic integration of fuzzy reasoning, the rough reasoning and case-based reasoning technology and multi-agent collaboration technology to achieve multi-inference engine joint qualitative forecast; effective integration of multiple knowledge fusion, multi-library collaborative knowledge discovery methods knowledge of related cases, experience, knowledge, and behavior models for virtual miner agent next step behavior of the decision-making guidance.

5 Realize the Behavior Simulation of the Virtual Miner Agent

This paper uses the MultiGen Creator real-time 3D modeling, using Vega or OpenGL real-time scene rendering. DI-Guy, the underlying library movement through the DI-Guy control interface, motion control virtual miner agent. DI-Guy is three-dimensional human body simulation software, virtual characters can be added to the real-time simulation environment, each virtual lifelike movement can
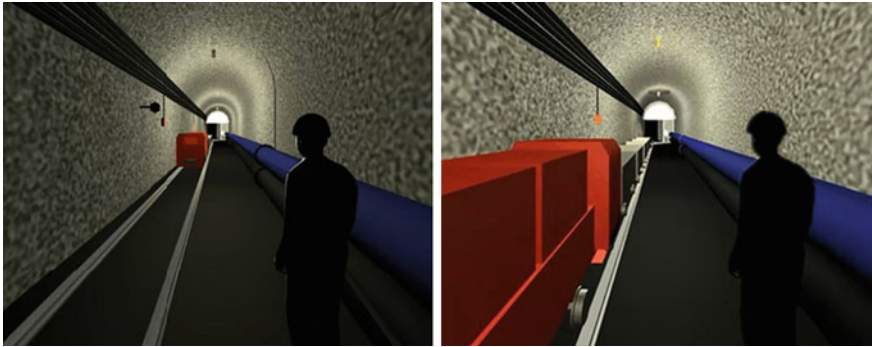
Fig. 3 Virtual miner circumvent the transport vehicle process

also respond to commands. DI-Guy can automatically drive each role, and the natural behavior from one human to another human behavior, realistic conversion can be completed. The virtual human is developed by DI-Guy can real-time interaction with the other agent in the virtual environment.

The virtual miners (the trainees' agent) in accordance with their capabilities, database knowledge, task planning and path planning, and advance to the target sites along the planned path. When virtual miner advancing will likely encounter steps, vehicles and other obstacles, it will be able to independently take obstacle avoidance behavior. Figure 3 is a virtual miner agent proceeds along the path planning process encountered in transport vehicles agent passage, before the encounter virtual miners agent will receive the transmission of the transport vehicle of the communication agent related information (transporter agent running speed and direction, and transport substances, etc.). At this time, the virtual miner will stand away from transport vehicles agent roadway edge forwardly, when other transport vehicles agent through before they continue to move forward towards the planned.

6 Conclusion

Based on the agent technology, constructing a virtual environment of the coal miner virtual agent system model and its formalization definition, realize the coal mine under virtual environment virtual miner agent of interactive control and underground safety behavior simulation. The simulation can effectively realize virtual miner's behavior modeling, not only can realize the path planning control, but also on the virtual environment in various events to make timely response. The results of simulation show that coal mine in the virtual environment of virtual miners agent can interact with trainer real-time to meet the preliminary needs of training.

References

1. Kasap Z, Magnenat-Thalmann N (2007) Intelligent virtual humans with autonomy and personality: state-of-the-art. Intell Decis Technol 1:3–15
2. Zhang J, Zhang T, Chen H (2011) Virtual reality technology and its application. Tsinghua University press, Beijing
3. Hao Z, Hao W, Yang G (2012) Virtual reality and its application in the coal mine. Commun Comput Inf Sci 206–213
4. Kijonka M, Kodym O (2012) Coal industry technologies simulation with virtual reality utilization. In: Proceedings of the 2012 13th international carpathian control conference, pp 278–283
5. Ye B, Bao N, Bai Z (2010) The study on building of virtual reality system in large surface coal mine. IFIP Adv Inf Commun Technol 346:173–178
6. Tang CL, Huang YR, Ling LY, Qu LG (2011) Design of mine virtual system based on OpenGL and it s realization. Saf Coal Mines 42(2):87–88
7. Li M, Chen J, Xiong W, Zhang P, Wu D (2008) VRLane: a desktop virtual safety management program for underground coal mine. In: Proceedings of SPIE-the international society for optical engineering, p 71432J
8. Yuan PY, Peng WL (2012) Research on the virtual simulation of emergency management for coal mining enterprises. Appl Mech Mater 121:4117–4121
9. Wang Q, Yang F (2012) Virtual mine 3D roadway modeling and its application. Int J Digit Content Technol Appl 6(8):89–97
10. Kijonka M, Kodym O (2012) Coal industry technologies simulation with virtual reality utilization. In: Proceedings of the 2012 13th international carpathian control conference, pp 278–283
11. Wooldridge M Work, Shi Chunyi and Zhang Wei and Xu Jinhui:Translation (2003) An introduction to multi-agent systems. Publishing House of electronics industry, Beijing
12. Sun Y (2011) Research on multi-agent based modeling methods for energy complex systems. East China University of Science and Technology, Shanghai

Research of Detection Method of Mine Motor Vehicle Pedestrian Based on Image Processing

Chaoli Tang, Lina Wang, Liguo Qu and Yourui Huang

Abstract This paper use infrared camera to collect the front image of motor cars, and to pretreat image based on genetic algorithm and normalized incomplete Beta function. Using pulse coupled neural network for image two value segmentation; and using improved fuzzy edge detection algorithm based on genetic algorithm for recognition the rail and using heuristic method for fitting the rails; once pedestrian recognition algorithm identified pedestrian, the alarm is immediately triggered. This system can efficiently identify the pedestrian near the track, judge and early warn the position of pedestrian; it is a new technology which can eliminate the potential safety hazard of motor vehicles in the transportational process.

Keywords Coal mine · Image processing · Pedestrian detection · Binary image

1 Introduction

Mine electric vehicle is mainly used for long-distance transportation of underground tunnel and ground, which completes the delivery of coal, gangue, material, equipment and personnel. Because of the complex environment of the underground coal mine, motor cars run frequently and long-distance. The roadway is narrow, which for

C. Tang · L. Wang · L. Qu · Y. Huang (✉)
Department of Electrical Engineering, Anhui University of Science and Technology,
Huainan 232001, Anhui, China
e-mail: hyr628@163.com

C. Tang
e-mail: chaolitang@163.com

L. Wang
e-mail: lnwang@aust.edu.cn

L. Qu
e-mail: qlg77@163.com

both motor vehicle roadway and miners walking channel, therefore, negligence of locomotive drivers, misjudge, paralysis, fatigue and other conditions, or miners inobservance of the rules of safe walking often appears. Locomotive drivers' safety is the key in motor vehicle driving, so the personal injury accident due to the above situation often appears. Therefore, it is necessary to put forward a new solution to handle the significant security risks in the process of motor transportation.

2 Pedestrian Monitoring Algorithm Based on Image Processing

2.1 Image Adaptive Correction

In the monitoring recognition system, the enhancement processing for the original image needs to be proceeded, which can improve the image visual effect, and is easy to judge and to identify image. But the intelligence and adaptive is poor in most traditional methods of image enhancement, it needs more manual intervention, which greatly limit its applied range. Therefore, establishing an image correction algorithm of adaptive adjustment which is based on image characteristics is worth expecting [1]. In the normalized incomplete Beta function $B(\alpha, \beta)$, different values of α, β correspond to different non-linear transformation curve. The paper uses genetic algorithm for non-linear transformation parameters α, β to mediate adaptively, and finally can achieve the image adaptive correction.

Generally, the determination of the values of α, β uses similar exhaustive method or the method of manual intervention, but the method wastes time and has no intelligence. The paper utilizes genetic algorithm to automatically find the optimal value of α, β in the non-linear transformation function. The steps are as follows:

1. Let $f(x, y)$ be the gray value of original image of coordinate (x, y), and $f'(x, y)$ is the gray value after computing. Normalized operating is carried out before the operation:

$$g(x, y) = [f(x, y) - L_{\min}] / [L_{\max} - L_{\min}]$$

2. Encoding with the use of genetic algorithm and generating the initial group, each chromosome contains two gene segments, which are respectively α, β. Each individual is corresponding to a nonlinear transformation $F(u), 0 \leq u \leq 1$, and the image gray is operated by using $F(u), 0 \leq u \leq 1$.

$$g'(x, y) = F[g(x, y)]$$

Fig. 1 Original image

3. The image output can got from $g'(x, y)$;
4. Using image quality evaluation function as the fitness function of the genetic algorithm:

$$Fitness(i) = \frac{1}{n} \sum_{x=1}^{M} \sum_{y=1}^{N} f^2(x, y) - \left[\frac{1}{n} \sum_{x=1}^{M} \sum_{y=1}^{N} f(x, y) \right]^2$$

In the above formula, M, N is the width and height of the image, $n = M \times N$, i representing a chromosome.

5. Selecting the biggest individual in the $Fitness(i)$ as the current generation of optimal individual, and updating other individuals, get the new group. Repeat the above steps until the optimal non-linear transformation function parameters α and β are got.

As shown in Fig. 1 is the original image, and Fig. 2 is the image after image adaptive correction based on the combined genetic algorithm with normalized incomplete Beta function. It can be seen that the image gray in Fig. 2 is wider than original image, meanwhile, the contrast changes obviously.

2.2 Image Binary Based on Pulse-Coupled Neural Network

The identity of the objects is based on the object's certain characteristics. A gray-scale image includes not only the interesting objects, but also the noise and

background. The interested may be just some of these areas, and the process of region extraction is called image segmentation [2]. In the paper, the region of the interested is the pedestrians who are on the front track. In order to extract the information of pedestrians out, may according to the difference between the gray-scale of figurine and background, setting the method of the gray threshold to identify the pedestrians on the front of the track, thus treating the figurine and background as the separation of binary images, the process is called binary image.

The commonly used method of binary, such as global threshold method, according to different targets on image segmentation using an optimal threshold, namely the threshold value used for each pixel in the image are equal, and threshold value has nothing with the position of part image. But in many cases, the gray value of the background is not a constant. The contrast of the object and the

Fig. 3 Value of two images
(pedestrian)

Fig. 4 Value of two images
(no pedestrian)
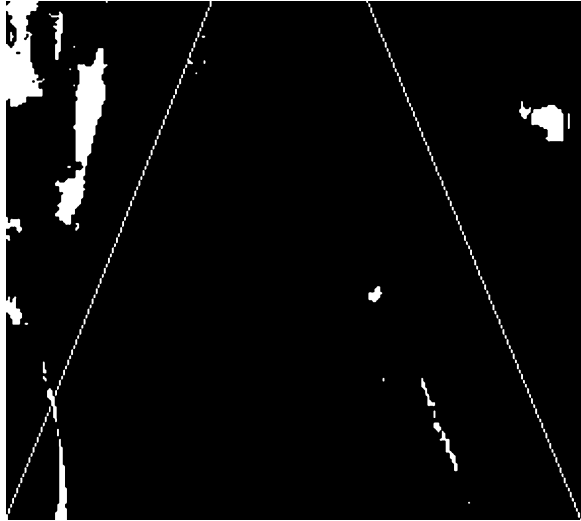
background in the image also changes. Here, a threshold value which is on a certain region in the image, its effect is good, but on other regions is not. So, in order to monitor and locate the pedestrians better, the paper puts forward the binary image based on the FPGA to achieve PCNN. PCNN has characteristics to group the pixel in the image of two-dimensional space similarity and gray-level similarity, and it can reduce the difference of the image local gray and make up the local small gap of the image. It has the incomparable advantages with other image segmentation methods [3–5].

Figures 3, 4 are the results of binary image in the track underground. The effect is relatively good.

3 The Identification and Fitting of the Track

The visual angle is much wider because of the pictures or videos collected by infrared video camera, and it includes much useless information. Therefore, the paper needs to identify and fit the picture tracks, which can reduce the processing region on one hand, on the other hand can accurately locate the position of the pedestrians. It mainly disposes the content about two aspects: road recognition of edge detection rails fitting. It has many algorithms for edge detection [6–8], which, Prewitt algorithm and Sobel algorithm have a good effect on edge detection, but the shortcoming is that the two are sensitive to the noise. The effect is good by using Robert algorithm to detect the horizontal and vertical edge, and the shortcoming is rough positioning. The advantage of LOG algorithm is isotropic. It not only detects the most edge, but also accurately locates the edges, and its disadvantage is the deletion of some key information of the image in the processing

results. The paper uses GFEDA to extract the edge which is the most integrated, and the continuity of the edge is good. Its effect is better than other algorithms. Figure 5 is the picture selected underground. By comparing with the above algorithms, proves GFEDA is superior.

The edge of the image in the original underground tracks has emerged clearly after the steps of image processing, edge extraction that have mentioned in the preamble. The paper uses heuristic connection methods to extract the two straight lines of the tracks edge, actually, that is to extract straight lines from images and further to judge the extraction whether is tracks' edge or not according to the prior knowledge.

Figure 6 is the experimental effect by using heuristic connection method to extract and fit the tracks under different conditions. Which, (a) is the straight tracks fitting with no noise, (b) is the straight tracks fitting with noise. In order to ensure the safety of the pedestrians, Fig. 6, the rails fitting is appropriate to increase the width between the two fitting curves.
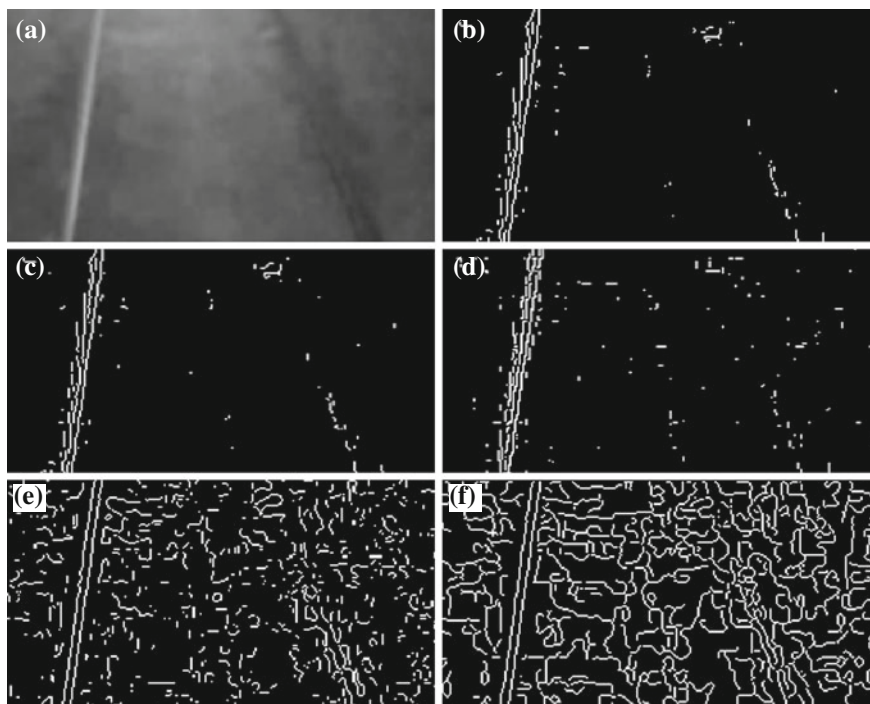
Fig. 5 Track boundary of different algorithms. **a** Image filtered. **b** Prewitt algorithm boundary extraction. **c** Sobel algorithm boundary extraction. **d** Roberts algorithm boundary extraction. **e** Log algorithm boundary extraction. **f** GFEDA algorithm boundary extraction
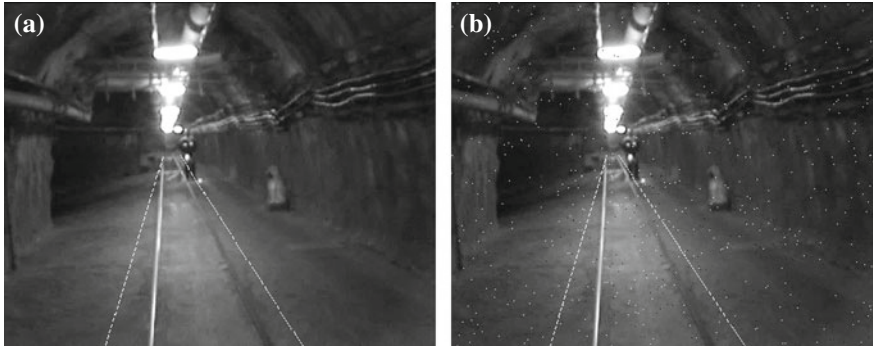
Fig. 6 The effect of Track fitting. **a** straight tracks fitting with no noise. **b** straight tracks fitting with noise

4 Identification of Pedestrians

The core idea of this article uses binary image method to detect the pedestrians effectively. First, segmenting figurine and background rapidly by adaptive background extraction method—it mainly combines the algorithm that the binary is based on the image of PCNN to the pedestrians segmented accurately. The method can effectively suppress the noise caused by the processing operation, and it can extract the complete contour of the pedestrians. Then, combining the feature information to preliminarily judge the existence of pedestrians, on the basis, locating and judging the pedestrians, get the accurate information of the pedestrians. The paper only needs to detect the pedestrians on the basis of the image preprocessing.

Figure 7 is the Pedestrian detection results of different distance. The results shows that the method used in this article can detect pedestrians who entry the image in real-time.
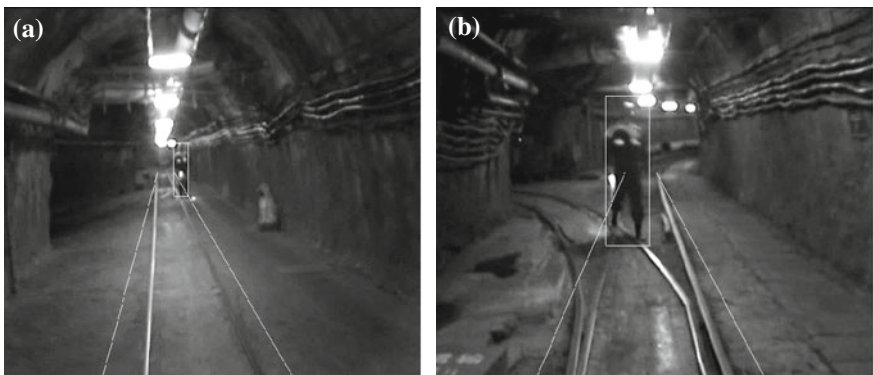
Fig. 7 Pedestrian detection results of different distance. **a** Remote testing results. **b** Close testing results

5 Conclusions

In this paper, the image processing technology is applied in the detection method of mine motor vehicle pedestrian. The article uses infrared video camera to collect the front image of motor cars, and processing the image by burning the internal high-speed DSP chip into the pedestrian recognition algorithm. It uses PCNN for binary image segmentation to decrease the difference of image local gray and to make up the image local small gap. Once the alarm system identifies the pedestrians, it immediately triggers the alarm. It can identify the pedestrians near the track effectively, and make judgments and early warning on its location. It is a new technology that can eliminate the hidden danger in the transportation of electric locomotive.

References

1. Dong L (2009) The study based on the FPGA digital image processing. Anhui University of Science and Technology
2. Yunzhi J, Hao Zhifeng, Yong LZ (2011) Automatic multi-threshold image segmentation based on Block sampling and genetic algorithm. J comput-aided des Comput Graph, 11:1860–1868
3. Dan song C, Liu X, Jianglong T etc. (2009) The image segmentation based on Neighborhood excitation pulse coupled neural network. J Huazhong Univ Sci Technol(Nat Sci), 5:33–37
4. Kong WW, Lei YJ, Lei Y, Lu S (2011) Image fusion technique based on non-subsampled contourlet transform and adaptive unit-fast-linking pulse-coupled neural network. IET Image Proc 2:113–121
5. Huang W, Jing Z (2007) Multi-focus image fusion using pulse coupled neural network. Pattern Recogn Lett 9:1123–1132
6. Libao Z, Dongling L, Xianchuan Y ect. (2010) Region of interest image edge detection based on histogram. High power laser and particle beams, 8:1847–1851
7. Bo Y, Luo S, Zou Q (2010) Natural image object contour significant computational model of the edge detection. Pattern Recogn Artif Intell 6:752–758
8. Duan R, Li Q, Li Y (2005) Summary of study on image edge detection. Opt Tech 3:415–419

Analysis of TBM Monitoring Data Based on Grey Theory and Neural Network

Tianrui Zhang, Yuanxing Dai, Caixiu Lu, Haifeng Zhao and Tianbiao Yu

Abstract This paper represented the main state parameters of Tunnel Boring Machine (TBM) system, analyzed the variation tendency of time series which were TBM characteristic parameters, predicted the development tendency for characteristic parameters of TBM equipment status combing the grey and neural network prediction, and then built the prediction model for characteristic parameters of TBM based on the grey theory and neural network. Through calculating the projects, the improvement measure of prediction model was given. The modified prediction model could ensure the running condition for 10 h when prediction accuracy reaches first class. Finally, this paper introduced the part of parameters prediction for the TBM fault diagnosis system developed by the author, so prediction results would be presented before the workers more directly.

Keywords TBM · Grey theory · Neural network · Monitoring data · Prediction analysis

1 Introduction

A lot of research had been done for developing the tunnel boring machine system. But until now, most evaluations for equipment status based on condition monitoring were from the view of security of operating the equipments. American EPRI declares the state evaluation module of operation and maintenance decision support system they developed had the ability of comprehensive evaluation [1–3]. The

T. Zhang (✉) · Y. Dai · C. Lu · T. Yu
School of Mechanical and Automation, Northeastern University, Shenyang, China
e-mail: tianjiangruixue@126.com

H. Zhao
TBM Company, Northern Heavy Industries Group Co., Ltd, Shenyang, China

integrated system for equipment maintenance of American Entek also included the module of comprehensive state evaluation. G. silnoes and TaehongKim developed a TBM usability prediction system using the fuzzy logic theory. S. Okubo, K. Fukui and W. Chen developed an expert system for evaluating the applicability of tunnel boring machine using in Japan [4, 5]. From the status of TBM fault diagnosis, although researches make progress, there wasn't a complete theoretical system or fault diagnosis system for tunnel boring machine [6].

This paper took the earth pressure balance shield machine as the research object, predicted the development tendency for characteristic parameters of TBM equipment status using the prediction module combing the grey theory and neural network theory, and built foundation for the complete and reasonable fault diagnosis system.

2 Type Style and Fonts Condition Monitoring Parameter Analysis of TBM

State parameters of TBM main system was shown as in Table 1.

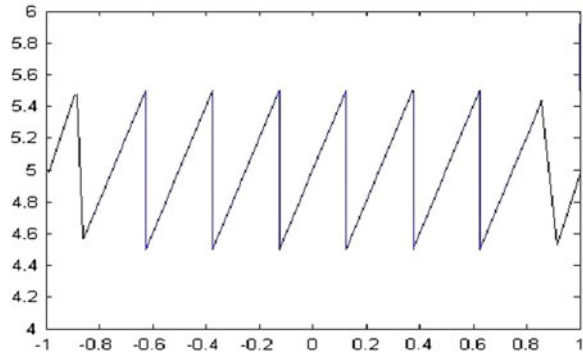### 2.1 Monitoring Parameter Prediction Analysis

Several common TBM characteristic parameters change trend as time sequence was shown as in Figs. 1, 2, 3, 4.

Due to the prediction model had its applicable condition, so different prediction model must be used for different types of characteristic parameters, then through the comprehensive evaluation to determine the state of the equipment in the future. Combined with the characteristics of TBM equipment state parameter and the
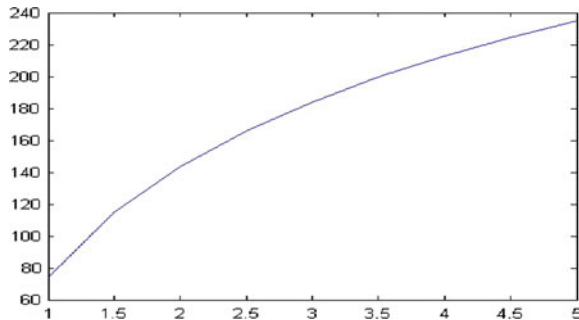
Table 1 Steady uptrend time series

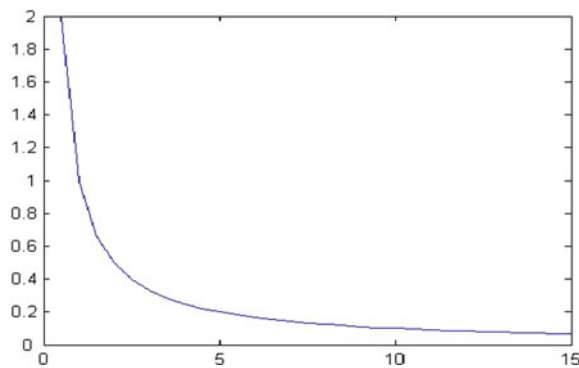| Systems	Parts	Monitoring parameter
Cutter drive system	Main bearing	Cutter speed
		…
		Oil pollution
	Reduction box	Input/output torque
		…
		Oil pollution
	Main motor	Current
		Voltage
Propelling system		…
Note fat system		…
…		…

Fig. 1 Stable fluctuate time series

Fig. 2 Steady uptrend time series

Fig. 3 Decline trend time series

applicable condition of prediction method, this paper had adopted grey prediction and neural network prediction to forecast the trend of characteristic parameters of TBM equipment state [7].
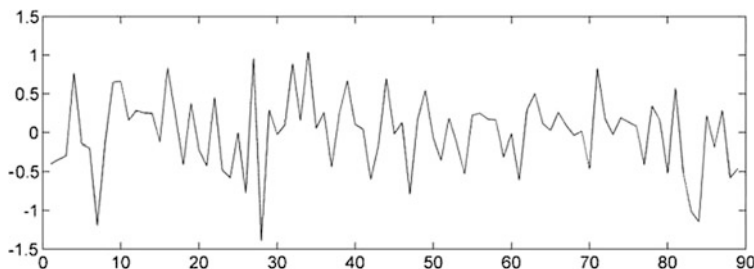
Fig. 4 Comprehensive trend time series

2.2 Established the Prediction Model of $GM(11)$, and Solution

It supposed that $X^{(0)}$ was the modeling sequence of $GM(11)$,

$$X^{(0)} = \left(x^{(0)}(1), {}^{(0)}(2), \ldots, {}^{(0)}(n) \right)$$

$X^{(1)}$ was the sequence of $X^{(0)}$ after one accumulator generated,

$$X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), \ldots, x^{(1)}(n) \right), x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), \ k = 1, 2, \ldots, n$$

Supposed $Z^{(1)}$ was the generated sequence of $X^{(1)}$ mean,

$$Z^{(1)} = (z^{(1)}(2), z^{(1)}(3), \ldots, z^{(1)}(n)), z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1)$$

Grey differential equation model of $GM(11)$, is

$$x^{(0)}(k) + az^{(1)}(k) = b \tag{1}$$

In formula (1), a was named development coefficient, b was the grey action. It supposed that $\hat{\alpha}$ is the waiting estimated parameter vector that was $\hat{\alpha} = (a, b)^T$, then the least squares estimate parameter list of grey equation should meet

$$\hat{\alpha} = \left(B^T B \right)^{-1} B^T Y_n \tag{2}$$

In formula (2),

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \ldots & \ldots \\ -z^{(1)}(n) & 1 \end{bmatrix}, \quad Y_n = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \ldots \\ x^{(0)}(n) \end{bmatrix}$$

Definition 1: $dx^{(1)}/dt + ax^{(1)} = b$ was the albino equation of grey differential equation $x^{(0)}(k) + az^{(1)}(k) = b$, also named shadow equation. As stated, the solution of $dx^{(1)}/dt + ax^{(1)} = b$ also called time response function,

$$\hat{x}^{(1)}(t) = \left(x^{(1)}(0) - b/a\right)e^{-at} + b/a \tag{3}$$

The response list of $x^{(0)}(k) + az^{(1)}(k) = b$ was

$$\hat{x}^{(1)}(k+1) = \left[x^{(1)}(0) - b/a\right]e^{-ak} + b/a, \ k = 1, 2, \ldots, n \tag{4}$$

If $x^{(1)}(0) = x^{(0)}(1)$, then

$$\hat{x}^{(1)}(k+1) = \left[x^{(0)}(1) - b/a\right]e^{-ak} + b/a, \ k = 1, 2, \ldots, n \tag{5}$$

Reducing value was

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \tag{6}$$

2.3 Established the Prediction Model of Artificial Neural Network and Solution

It supposed that has obtained $n+1$ time list samples $x(0), x(1), \ldots, x(n)$ of a certain state characteristic parameters, now the first n observed value were used to forecast the value $x(n+1)$ of $(n+1)th$ moment. The steps of learning and prediction used BP network shown as Fig. 5.

Step 1: Determine the number of network layer and unit number each layer. Got m units from input layer and one unit from output layer, but the hidden layer was one layer generally.

Step 2: Samples $x(0), x(1), \cdots, x(n)$ were divided into groups, and each group had $m + 1$ values, the first m values were input values and the $(m + 1)th$ value was output value. As shown in Table 2.

Step 3: BP arithmetic used to train network, and determined the connection weights of the network.

Step 4: Time series of condition characteristic parameters $x(n - m + 1)$, $x(n - m + 2), \ldots, x(n)$ were as network input, used the after trained network to forecast, the network output was the forecast value $x(n + 1)$ of $(n + 1)th$ moment [8].
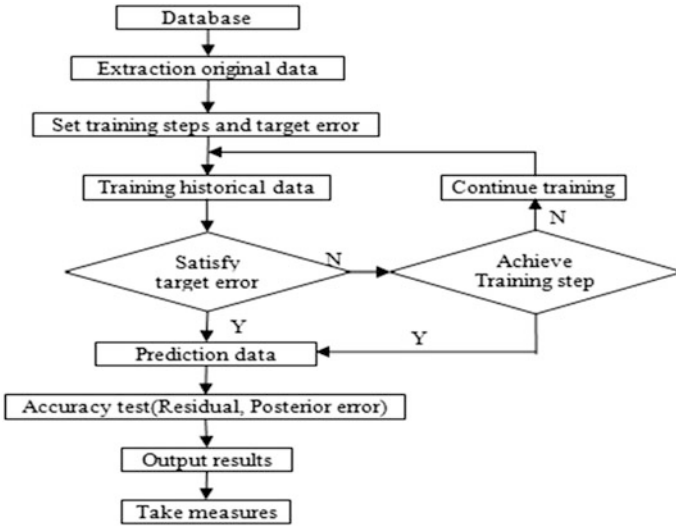
Fig. 5 Flow chart of prediction by neural network

Table 2 Input and output training samples of the BP neural network

Input	Output
$x(0), x(1), \ldots, x(m-1)$	$x(m)$
$x(0), x(1), \ldots, x(m)$	$x(m+1)$
…… …	…… …
$x(n-m), x(n-m+1), \ldots, x(n-1)$	$x(n)$

3 Model Test

Prediction model tests contained residual test, correlation test and posterior error test [9].

1. Residual test: calculating the absolute error and relative error between the original data and predicted data.

   Absolute error: $\varepsilon^{(0)}(i) = x^{(0)}(i) - \hat{x}^{(0)}(i), (i = 1, 2, \ldots, n)$.

   Relative error: $\omega^{(0)}(i) = \left| [x^{(0)}(i) - \hat{x}^{(0)}(i)]/x^{(0)}(i) \right|, (k = 1, 2, \ldots, n)$

   Here into, $\hat{x}^{(0)}(i) = \hat{x}^{(1)}(i) - \hat{x}^{(1)}(i-1), (i = 1, 2, \ldots, n)$, the relative error was smaller, the higher the accuracy of model.

2. Posterior error test: calculating the mean square deviation of original sequence $x^{(0)}(i)$ firstly:

$$S_0 = \sqrt{S_0^2/n - 1}, \text{ while } S_0^2 = \sum_{i=1}^{n} \left[x^{(0)}(i) - \bar{x}^{(0)} \right]^2, \bar{x}^{(0)} = \frac{1}{n}\sum_{i=1}^{n} x^{(0)}(i).$$

Table 3 Standard table of forecast accuracy class

Small error probability p	Variance ratio c	Forecast accuracy grade
>0.95	<0.35	Grade 1 (excellent)
>0.80	<0.50	Grade 2 (favorable)
>0.70	<0.65	Grade 3 (qualification)
≤0.70	≥0.65	Grade 4 (disqualification)

Then calculating the mean square deviation of residual error sequence $\varepsilon^{(0)}(i)$:

$S_1 = \sqrt{S_1^2/n-1}$, while $S_1^2 = \sum_{i=1}^{n} \left[\varepsilon^{(0)}(i) - \bar{\varepsilon}^{(0)}\right]^2, \bar{\varepsilon}^{(0)} = \frac{1}{n}\sum_{i=1}^{n}\varepsilon^{(0)}(i)$.

$$\text{Variance ratio, } c = S_1/S_0 \tag{7}$$

Small error probability

$$p = \left\{ \left| \varepsilon^{(0)} - \bar{\varepsilon}^{(0)} \right| < 0.6745 \cdot S_0 \right\} \tag{8}$$

According to the forecast accuracy grade classification to determine the accuracy of model that shown as Table 3 [10].

4 Application Examples of Prediction Model

Hydraulic oil's temperature changing trend was growth smoothly, it was suit to use $GM(11)$, model to fitting and forecast. The change trend of the main bearing's vibration amplitude was bigger, so BP neural network only could be used to fitting and forecasting. The experiment were taken with the 30 groups historical data (Data extraction frequency was 2 h) of two feature state parameters, while the top

Fig. 6 Forecasting curve of main bearing-vibration breadth based on neural networks
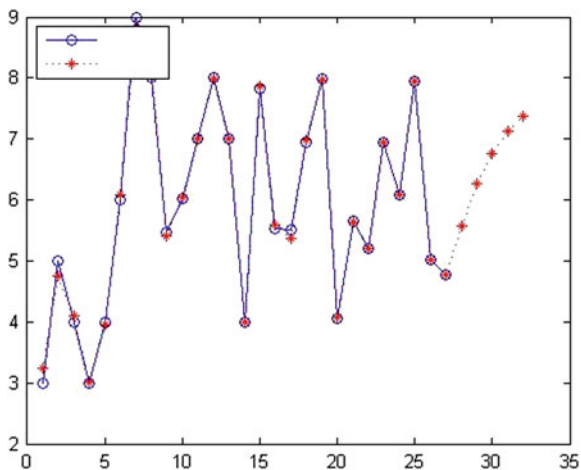
Fig. 7 Absolute error of
vibration breadth
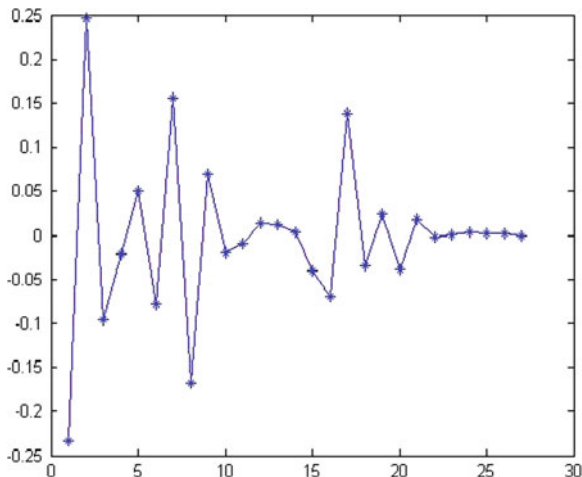
25 groups data were as the training sample, the last 5 groups data were used for model test. The results could be seen in Figs. 6, 7 and Table 4.

Table 4 showed that the maximum forecasting 5 steps (10 h) of the state parameters could be realized by using these two kinds of methods in the scope of allowable error. But the prediction error was increasing with the prediction step number. Improved measures: grey forecasting could be corrected by using residual model; neural network prediction could be corrected by increasing the network output unit or improving the algorithm amendment. Because of the real-time monitoring of TBM, this model could predict the condition of the equipment operation for 10 h in first level forecasting precision; this has met the application requirements.

Table 4 Forecast results of state characteristic parameters

Detected feature	Measure value	Predicted value	Relative error (%)	Small error probability p
Hydraulic oil temperature	35	35.6	0.0029	$p = 1.0$ prediction accuracy was grade 1
	38	37.9	0.0031	
	37	36.6	0.0040	
	40	38.9	0.0039	
	38	37.56	0.0053	
Main bearing amplitude	7.137	6.9516	0.0069	$p = 0.096$ prediction accuracy was grade 1
	5.985	5.9766	0.0039	
	7.982	7.8884	0.0140	
	5.105	5.0597	0.0119	

5 Operation Examples of System

The real-time fault diagnosis expert system integrated all function modules of the TBM real-time fault diagnosis expert system, which included data mining, case-based reasoning, rule-based reasoning, database maintenance, system mainte-nance, help and so on.

This paper mainly introduced the part of the data analysis. These functions mainly included data extracting, browsing and Forecasting of inspection data, data discrimination and feature weight calculation.

This system has the function of prediction analysis of real-time detector data, it can view and prediction analysis of needed data through the database in advance when it is necessary. Choose the detection index in pull-down menu and examine, Then the real data will display in the list. If you want to view the future trend of index, you can choose prediction method, fill in the prediction step and click the prediction analysis button. So it will be predicted by the steps which you filled in, and display the prediction data and curves of real data which varying with time, it is easy for the technicists to make the decision.

6 Conclusion

1. $GM(11)$, and neural network prediction model was used in this paper, it could prediction analysis all kinds of time series. In addition, it could foreknow the operation status of TBM devices, which could lessen the accidents.
2. Because the prediction error increased with prediction step, it was improved in the paper. With the improvement, it could improve predicting accuracy and made the device reach first level forecast accuracy, and predicted the operation status of device 10 h later.
3. This paper only introduced the data analysis and pretreatment of TBM fault diagnosis system. To guarantee the integrality of the study, the author would introduce fault diagnosis model and technical service model in the future.

References

1. Sutherland PE (2010) Performance calculation for tunnel boring machine motors. Industry applications conference, conference record of the 2010 IEEE, vol 4, pp 2668–2673
2. Simoes MG, Kim T (2010) Fuzzy modeling approaches for the prediction of machine utilization in hard rock tunnel boring machines. Industry applications conference, conference record of the 2010 IEEE, vol 2, pp 947–954
3. Grima MA, Bruines PA, Verhoef PNW (2009) Modeling tunnel boring machine performance by neuro-fuzzy methods. Tunneling Underground Space Technol 15:259–269

4. Okubo S, Fukui K, Chen W (2011) Expert system for applicability of tunnel boring machines in Japan. Rock Mech Rock Eng 36:305–322
5. Skowron (2006) Extracting laws from decision tables: a rough set approach. Comput Intell 2:371–388
6. Holland (1962) Concerning efficient adaptive systems. Spartan Books, Washington, pp 215–230
7. Holland (2005) Adaption in natural and artificial systems. University of Michigan Press, Ann Arbor
8. Watson I (1997) Applying case-based reasoning techniques for enterprise systems. Morgan Kaufmann Publishers, Burlington
9. Watson I (1994) Case-based reasoning: a review. Knowl Eng Rev 9:335–381
10. Cser L (1991) Three kinds of case-based learning sheet manufacturing. Comput Ind 17:195–206

Kernel P Systems: Applications and Implementations

**Florentin Ipate, Raluca Lefticaru, Laurenţiu Mierlă,
Luis Valencia Cabrera, Huang Han, Gexiang Zhang,
Ciprian Dragomir, Mario J. Pérez Jiménez and Marian Gheorghe**

Abstract This paper explores the modelling capacities of a new class of P systems, called kernel P systems (kP systems). A specific language for describing kP systems and its translation into Promela, the specification language of Spin, are described.

F. Ipate (✉) · R. Lefticaru
Faculty of Mathematics and Computer Science, University of Bucharest, Academiei 14,
Bucharest, Romania
e-mail: florentin.ipate@ifsoft.ro

R. Lefticaru
e-mail: raluca.lefticaru@gmail.com

F. Ipate · R. Lefticaru · L. Mierlă · M. Gheorghe
Faculty of Mathematics and Computer Science, University of Piteşti, Târgu din Vale 1,
Piteşti, Romania
e-mail: laurentiu.mierla@gmail.com

M. Gheorghe
e-mail: m.gheorghe@sheffield.ac.uk

L. V. Cabrera · M. J. P. Jiménez
Research Group on Natural Computing, Department of Computer Science and Artificial
Intelligence, University of Sevilla Avda, Reina Mercedes s/n 41012 Sevilla, Spain
e-mail: lvalencia@us.es

M. J. P. Jiménez
e-mail: marper@us.es

H. Han · G. Zhang
School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, People's
Republic of China
e-mail: 329298217@qq.com

G. Zhang
e-mail: zhgxdylan@126.com

C. Dragomir · M. Gheorghe
Department of Computer Science, University of Sheffield, Portobello Street, Regent Court,
Sheffield S1 4DP, UK
e-mail: c.dragomir@sheffield.ac.uk

This Promela specification has been further used for simulation and property verification with the Spin model checker. Also, a parallel implementation on GPU parallel architectures, realized using CUDA, is presented and the results are compared with the ones obtained using Promela and Spin. A case study, namely the Subset sum problem, which has been modelled with kernel P systems and further implemented in Promela is presented.
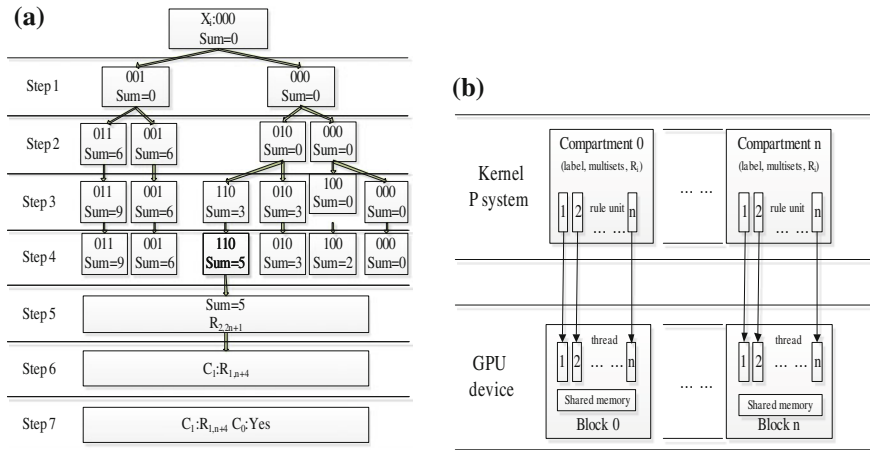
1 Introduction

Membrane computing, initiated in the seminal work of Păun [1], is a new branch of natural computing, whose models, *membrane systems* or *P systems*, are inspired from the structure and the functioning of the living cell and of other more complex systems, like tissues, organs or populations of cells [2]. In the last years, many P system models have been proposed and important developments have been reported [3], concerning the computational power of different variants, their capabilities to solve NP-complete problems, decidability or complexity aspects. Furthermore, these distributed parallel computing models have been successfully used to model, simulate and formally verify various systems [3, 4].

Kernel P systems (*kP systems*, for short) represent a new class of membrane systems, recently proposed in [5]. They use well-known features of P systems and also include new elements, in order to provide a coherent and comprehensive description of various classes of P systems and allow the possibility of integrating different types of P systems into the same formalism. For example, neural-like P systems and P systems with active membranes have been simulated by the newly introduced kP systems [5]. Other initial case studies, applying kP systems to solve the 3-colouring problem, previously solved with tissue P systems, have been presented in [6, 7] and analysed in terms of efficiency and complexity.

The expressive power of kP systems is further investigated in this paper and their modelling capacity is illustrated on a new case study, representing another NP-complete problem. Furthermore, compared with [6], additional modalities to implement kP systems are given and the model checker Spin [8] has been used to simulate and verify properties expressed for the Promela model. A parallel implementation of P systems using GPU and CUDA is also presented, and some initial results are reported.

Fig. 1 **a** The procedure of the kernel P system for solving the subset sum for $n = 3$. **b** Basic design of the simulation

2 Background

In this section we formally introduce a particular type of *kernel P systems*, namely *simple kernel P systems*, or *skP systems* for short [6]; a description of the general *kP system* model is provided in [5]. The skP systems have a graph-like structure, similar to *tissue P systems*, *compartments*, *multisets of objects* in them and *rules* of two types: (a) *object processing* and (b) *system structure* rules. The kP system rules have *guards*, with activators and inhibitors, described in a more general way than in traditional P systems. The execution strategy applied in such systems will be *maximal parallelism*, traditionally used in membrane computing, but, in the context of general kP system formalism, other types of strategies can be defined.

In the rest of the paper we use the following notations. For an *alphabet A*, the set of all strings over A is denoted by A^*. When the empty string, λ, is not considered, then the set is denoted by A^+. If u and v are strings over A, then their *concatenation* is given by uv. The *length* of the string u is denoted by $|u|$. If $u \in A^*$ and $a \in A$ then $|u|_a$ denote the number of occurrences of symbol a in string u. A *multiset m* over a set A is a pair (A, f) where $f : A \rightarrow \mathbf{N}$ is a mapping. The *support* of the multiset m is defined as $supp(m) = \{x \in A | f(x) > 0\}$. If $m = (A, f)$ is a finite multiset over A and $supp(m) = \{a_1, \ldots a_k\}$, then it will be denoted as $m = \left\{ a_1^{f(a_1)}, \ldots, a_k^{f(a_k)} \right\}$ or represented by a string $a_1^{f(a_1)} \ldots a_k^{f(a_k)}$.

Definition 1 Let $Rel = \{<, \leq, \neq, \geq, >\}$ be a set of relational operators. Let A be a nonempty finite set and denote $\bar{A} = \{\bar{a} | a \in A\}$. A *multiset over $A \cup \bar{A}$ with relational operators from Rel* is an expression $w = \theta_1 a_1^{n_1} \ldots \theta_k a_k^{n_k}$, where $a_1^{n_1} \ldots a_k^{n_k}$ is a string over $A \cup \bar{A}$ (in which a_i and a_j are not necessarily distinct, $1 \leq i < j \leq k$), and $\theta_j \in Rel$, for each j, $1 \leq j \leq k$.

Definition 2 A *guard* g is a finite disjunction of expressions w introduced by Definition 1.

A particular case of a guard g is the empty set. In this case the guard is omitted. The elements of the set A are considered activators and those of $\bar{A}$ inhibitors. For example, a guard $g = {<}a^3 = \bar{b}$ is true if the current multiset contains at most two objects a and no b, for instance a^2c.

Definition 3 Given two non empty finite sets A (alphabet), L (labels), and $l \in L$, the set of rules associated with l is denoted by R_l. A rule from R_l has one of the following two forms: (**a**) $[x]_l \rightarrow [y_1]_{l_1} \ldots [y_h]_{l_h} \{g\}$, where $x \in A, y_j \in A^*, l_j \in L$ for all $1 \leq j \leq h$, and g is a guard (*h-membrane division* rule); (**b**) $x \rightarrow y\{g\}$, where $x \in A^+$, y is a string over $A \times L$, $y = (a_1, t_1) \ldots (a_h, t_h)$, with $a_j \in A, t_j \in L$ and g is a guard (*rewriting and communication* rule).

Definition 4 Given two nonempty finite sets A, L, a *compartment*, C, is a tuple (l, w_0, R_l), where $l \in L, w_0$ is a multiset over A, and R_l is a set of rules associated with l.

A compartment can be viewed as a region labelled by an element l of L which initially contains a multiset w_0 of objects (elements of A) and such that it has associated a set R_l of rules.

Definition 5 Let r be a rule from R_l, a set of rules associated with a compartment $C = (l, w_0, R_l)$, such that the guard of r is g. The guard g is considered *true* at an instant when the current multiset of compartment C is z if and only if the following happens: (**a**) if $g = \theta_1 a_1^{n_1} \ldots \theta_k a_k^{n_k}$, then for every j, $1 \leq j \leq k$, either (i) if $a_j \in A$ then $|z|_{a_j} \theta_j n_j$ holds, or (ii) if $a_j \in \bar{A}$, then $|z|_{a_j} \theta_j n_j$ does not hold; (**b**) if g is a finite nonempty disjunction of multisets over $A \cup \bar{A}$ with relational operators from *Rel*, $g = w_1 | \ldots | w_p$, then there exists j, $1 \leq j \leq p$, such that w_j is true, according to (*a*); (**c**) if g is the empty set then it is always evaluated true.

Definition 6 A *simple kernel P* (*skP*) system of degree $n \geq 1$ is a tuple, $sk\Pi = (A, L, IO, C_1, \ldots, C_n, \mu, i_0)$ where: A and L are nonempty finite sets; IO is a finite alphabet $IO \subseteq A$; $C_1, \ldots, C_n$ are compartments; $\mu = (V, E)$ is an undirected graph, where $V \subseteq L$ are vertices and E the edges; $i_0 \in L$.

A skP system, $sk\Pi = (A, L, IO, C_1, \ldots, C_n, \mu, i_0)$, can be viewed as a set of n compartments, $C_1, \ldots, C_n$, interconnected by edges from E, of an undirected graph μ. The elements of the set A are called *objects* and the elements of L are called *labels*. IO is the alphabet of the *environment objects*. Each compartment is identified by a label of L, has initially a multiset over A, and a finite set of rules. The compartment receiving the result of a computation is denoted by i_0; in the sequel this will always be the *environment*.

A *h-membrane division* rule $[x]_l \rightarrow [y_1]_{l_1} \ldots [y_h]_{l_h} \{g\}$ associated with a compartment $C = (l, w_0, R_l)$ is applicable at a given instant to the current multiset z, if the guard g is evaluated true with respect to z and the object x is in the multiset z. When applying such a rule to x, the compartment labelled l will be replaced by

h compartments labelled $l_1,...,l_h$ and x is replaced by multiset y_j in compartment l_j; the content of l, but x, after using all the applicable rewriting and communication rules is copied in each of these compartments; all the links of l are inherited by each of the newly created compartments. A rewriting and communication rule $x \rightarrow (a_1, t_1)...(a_h, t_h)\{g\}$ associated with a set of rules, R_l, of a compartment $C = (l, w_0, R_l)$, is applicable at a given moment to the current multiset z, if the guard g is evaluated true, x is contained in z and the target $t_j \in L$, $1 \leq j \leq h$, must be either the label of the current compartment, l, or of an existing neighbour of it $((l, t_j) \in E)$. When applying the rule, objects a_j are sent to the compartment labelled by t_j, for each j, $1 \leq j \leq h$. If a target, t_j, refers to a label that appears more than once then one of the involved compartments will be non-deterministically chosen. When t_j indicates the label of the environment, then a_j is sent to the environment.

3 Modelling with Kernel P Systems

In order to illustrate the modelling and expressive power of skP systems, as well as their efficiency, the *Subset Sum* problem [9, 10] has been considered. This problem can be formulated as follows: given a finite set V, a function *weight* on V with positive integer values (an additive function), a number k representing the target sum, decide if there exists a subset W of V, such that $weight(W) = k$.

Theorem 1 *The Subset Sum problem for a set with* n *elements can be solved by a skP system and an answer to whether a solution exists or not is obtained in at most* n + 3 *steps using maximum* 2^n + 1 *compartments.*

Proof A sketch of the proof is provided below. For a given set $V = \{v_1,..., v_n\}$, with $weight(v_i) = k_i$, where k_i is a positive integer, $1 \leq i \leq n$, we build the following skP system, which depends on n, (for checking whether there is a subset, W, with $weight(W) = k$), $sk\Pi(n) = (A, L, IO, \mu, C_1, C_2, 0)$ where

- $A = \{A_1,..., A_n\} \cup \{B_1,..., B_n\} \cup \{v_1,..., v_n\} \cup \{v, S, T, F, X, Y, yes, no\}$
  where A_i, $1 \leq i \leq n$ are used in the division process of C_2 compartments, until the X element is used to control the end of the process; B_i, $1 \leq i \leq n$ stand for the n elements in set; v_i, $1 \leq i \leq n$ are used to represent the weight of each element, and v the total weight for each compartment; Y is used to trigger the evaluation of weight v against k, generating T (true) or F (false); *yes*, *no* are the two possible answers depending on the objects T, F;
- $L = \{0, 1, 2\}$; 0 is the label of the environment, 1 and 2 of the compartments;
- IO consists of *yes*, *no*; after $n + 3$ steps, one of them will be sent out;
- $C_1 = (1, w_{1,0}, R_1)$, $C_2 = (2, w_{2,0}, R_2)$, where $w_{1,0} = S$, $w_{2,0} = A_1 code(n)$, with $code(n) = v_1^{k_1}...v_n^{k_n}$ being the code of the weights of the elements of V;
- μ is given by the graph with edge $(1, 2)$;

- R_1 contains: $r_{1,1} : S \rightarrow (yes, 0)\{ \geq T \}, r_{1,2} : S \rightarrow (no, 0)\{ \geq F \geq \overline{T} \}$; $r_{1,\,1}$ or $r_{1,\,2}$ sends into the environment the answer *yes* or *no*, respectively;
- R_2 contains: *membrane division rules* $r_{2,i} : [A_i]_2 \rightarrow [B_i A_{i+1}]_2 [A_{i+1}]_2$, $1 \leq i < n, r_{2,n} : [A_n]_2 \rightarrow [B_n X]_2 [X]_2$; these rules generate in n steps all the subsets of V (2^n subsets); each of them being a potential subset W; *rewriting rules* $r_{2,n+i} : v_i \rightarrow v\{= B_i = X\}, 1 \leq i \leq n, r_{2,2n+1} : X \rightarrow Y$ and *rewriting and communication rules* $r_{2,2n+2} : Y \rightarrow (F, 1)\{ \neq v^k \}$, $r_{2,2n+3} : Y \rightarrow (T, 1)\{ = v^k \}$.

The computation leads to an answer, *yes* or *no*, in $n + 3$ steps.

4 Experimental Evaluation

This section is devoted to skP system simulation and verification of certain properties using the Spin model checker [8]. The skP system is converted into Promela, the specification language for Spin. This is obtained automatically from the skP system specification. A simple specification language allows to express skP system elements and declarations in a simple, concise and intuitive way. The complete implementation of the Subset Sum problem as modelled with kP systems is provided at [11]. Spin [8] is a very popular simulator and model checker tool, purposely built for the formal verification of distributed software systems.

We will refer to an instance of the Subset Sum problem where the number of elements in our set is 10 (the example is available at [11]). Given this parameter, we define a skP system, according to the algorithm illustrated in the previous section, that requires two active compartments, an alphabet of 38 symbols and a total of 25 rules. A successful simulation requires 13 computational steps, the outcome of which is a "*yes*" symbol that is sent to the environment, indicating that there is at least one subset with the sum of its elements equal to the specified value. In the context of membrane computing, Spin has been successfully used to formally verify various types of P systems [12, 13]. The properties we investigate are expressed as LTL formulae and their validity is assessed with respect to our model, an instance of the Subset Sum problem having 10 set elements. The key insight of our approach is the distinction between a P system step and a generic step of a Promela model, which takes into account each atomic instruction. However, a P system's state is reached only after the maximally parallel application of the rules in every compartment. Hence, an auxiliary variable is introduced, indicative of whether the system has reached a state or it is in the process of executing rules. This strategic flag is then mentioned in the LTL properties, instructing Spin to only consider states with relevant configurations (i.e. P system states) in its evaluation. A common property pattern in our study was that of necessary precedence: if a certain condition *invariantPrecond* holds for a state, then another condition *invariantPostcond* must hold in the next state. This is expressed in LTL as: *ltl p5 {[] (!(invariantPrecond) || X(!isPsystemStep U(invariantPostcond && isPsystemStep)) || !isPsystemStep)},*

where *invariantPrecond* and *invariantPostcond* are the two subject conditions and *isPsystemStep* is a Boolean variable which acts as a filter on states, as we are only interested if the property holds for a P system state. A concrete formula that exemplifies this pattern is *ltl p4 {[] (!(ms[1].x[@S] == 1 && ms[1].x[@F] == 1) || X(!isPsystemStep U (ms[2].x[@no] == 1 && isPsystemStep)) || !isPsystemStep)}*, which asserts that if in a certain state the membrane of type "1" contains both an *S* and an *F* object, then, in the next state, the environment will have received a "*no*" object. Another interesting property successfully validated states that if a *Y* symbol has been generated into the compartments of type 2, then, at the next P system step, the multiplicity of the object *T* in the compartment of type 1 will be equal to the number of solutions of the problem. The property is captured by the following formula: *ltl p6 {[] (!(invariantPrecond) || X(!isPsystemStep U(solution-Count == TCount && isPsystemStep)) || !isPsystemStep)}*.

5 Implementation of kP Systems Using GPUs

In order to solve the Subset Sum problem (SSP) using GPUs, we make the following notations. We start with a membrane structure $\mu = [[]_1[]_2]_0$, compartment labels $L = \{0, 1, 2\}$ and symbols $A_1, \ldots, A_{n+4}, B_1, \ldots, B_n, v_1, \ldots, v_n, T, Y, K, V, yes, no$, where A_i stands for the *i*th step; v_i represents the *i* element of the set; B_i means whether v_i is in the subset or not; object K is the desired subset sum; V is a generic element counting for the weight of the subset; *Y* is an object which would be changed into a flag and be sent to the compartment C_1 to show the objects in this compartment are a solution and *T* is a flag just mentioned; *yes* and *no* are the two possible answers sent from C_1 to the environment (C_0) at the end of the computation. The compartments are $C_0 = \emptyset$, $C_1 = (1, O_1, R_1)$, $C_2 = (2, O_2, R_2)$, where $O_1 = \{A_1\}, O_2 = \{A_1, v_1^{k_1} \ldots v_n^{k_n}, Y\}$, where v_i is the value of subset *ith* element and k_i is its weight. All the rules are executed in a maximally parallel way and are described as follows. It is worth pointing out that there are some differences between the rules here and the rules in Theorem 1. The rules considered here are used to stop the execution in some compartments when the current value inside the compartment is bigger than the desired subset sum, thus the execution time and memory required can be saved to a certain degree. R_1 contains $r_{1,i} : A_i \rightarrow A_{i+1}, 0 < i < n + 4; r_{1,n+4} : T \rightarrow (yes, 0)\{A_{n+4}\}; r_{1,n+5} : A_{n+4} \rightarrow (no, 0)\{\overline{T}\}$; R_2 contains $r_{2,i} : [A_i]_2 \rightarrow [A_{i+1}, B_i]_2[A_{i+1}]_2\{<V^k\}, 0 < i < n + 1; r_{2,n+i} : v_i \rightarrow V\{= B_i\}, 0 < i < n + 1; r_{2,2n+1} : Y \rightarrow (T, 1)\{= V^K\}$.

Example Figure 1a shows the procedure for solving the Subset Sum problem for 3 elements with weights 2, 3, 6 and $K = 5$. The compartments on the left side in Steps 2–4 could only execute $r_{2, n+1}$ for the guards of the rest rules, which requires the Sum is smaller than or equal to K, in compartment C_2.
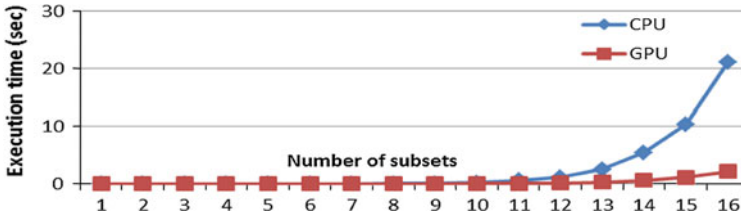
Fig. 2 Comparison of the elapsed time

Figure 1b shows the basic design of the simulation. We define the GPU blocks as the compartments and the GPU threads as units which execute the rules. Each thread and block runs in parallel. All the parameters in each compartment are sent to the GPU global memory. To decrease the time for reading data, we store constants and rules, which are never changed, into the constant memory or texture memory. For the same reason, the compartment parameters are sent to the shared memory in blocks, when the compartment chooses and applies the rules. Because of the restriction of GPU on the number of blocks and threads, the number of the compartments and rules should be less than 65536 and 512, respectively.

This simulation is performed in two stages: selection and execution. The former stage is to search for the rules to be executed in each compartment. The latter stage is used for performing the rules selected at the previous stage. According to the kernel P system definition, the selection stage includes the processes of checking the guard and judging whether one rule is able to be performed. At the execution stage, the compartment can receive new objects, rewriting objects, transition objects. The membrane structure may be changed due to division.

The platform environment is as follows: CPU i5-3240M 2.50 GHz, Memory 8 GB DDR3, NVIDIAGeForce GT650M 1 GB DDR3, Operating system: Windows 7 Professional. The experimental results are shown in Fig. 2 (The number of the use of $r_{2, n+i}$ will be very large with the increase of the value of k_i, so the value of each weight, in our simulation, is less than 200).

It can be seen from the results shown in Fig. 2 that as the number of subsets increases, the difference of the elapsed time between GPU and CPU becomes larger, and that the elapsed time on GPU is about 10 times faster than that on CPU.

6 Conclusions

This paper illustrates the modelling capabilities of the new class of kP systems on solving a well-known NP-complete problem, namely the Subset Sum. For this model, a kP system solution is provided and a Promela executable specification has been written and used with the Spin model checker for simulation and property verification. A parallel implementation on GPUs is presented and the results obtained show the efficiency of this implementation wrt CPU one.

References

1. Păun G (2000) Computing with membranes. J Comput Syst Sci 61:108–143
2. Păun G (2002) Membrane computing: an introduction. Springer, London
3. Păun G, Rozenberg G, Salomaa A (eds) (2010) The Oxford handbook of membrane computing. Oxford University Press, Oxford
4. Ciobanu G, Pérez-Jiménez MJ, Păun G (eds) (2006) Applications of membrane computing., Natural computing seriesSpringer, London
5. Gheorghe M, Ipate F, Dragomir C (2012) A kernel P system. In: Proceedings of 10th brainstorming week on membrane computing. Fénix Editora, Seville, Spain, pp 153–170
6. Gheorghe M, Ipate F, Lefticaru R, Pérez-Jiménez MJ, Ţurcanu A, Valencia Cabrera L, García-Quimondo M, Mierlă L (2012) 3-col problem modelling using simple kernel P systems. Int J Comput Math. doi:10.1080/00207160.2012.743712
7. Ipate F, Dragomir, C, Lefticaru R, Mierlă L, Pérez-Jiménez MJ (2012) Using a kernel P system to solve the 3-col problem. In: Proceedings of the 13th international conference on membrane computing. Computer and Automation Research Institute, Hungarian Academy of Sciences, pp 243–258
8. Ben-Ari M (2008) Principles of the spin model checker. Springer, London
9. Leporati A, Mauri G, Zandron C, Păun G, Pérez-Jiménez MJ (2009) Uniform solutions to SAT and subset Sum by spiking neural P systems. Nat Comput 8:681–702
10. Pérez-Jiménez MJ, Riscos-Núñez A (2005) Solving the subset-Sum problem by P systems with active membranes. New Gener Comput 23(4):339–356
11. Code repository http://www.p-lingua.org/mecosim/doc/case_studies/SubsetSumBIC-TA/
12. Ipate F, Lefticaru R, Pérez-Hurtado I, Pérez-Jiménez MJ, Tudose C (2012) Formal verification of P systems with active membranes through model checking. In: Gheorghe M, Păun G, Rozenberg G, Salomaa A, Verlan S (eds) Proceeding of the international conference on membrane computing 2011. Lecture Notes in Computer Science, vol 7184. Springer pp 215–225
13. Ipate F, Lefticaru R, Tudose C (2011) Formal verification of P systems using Spin. Int J Found Comput Sci 22:133–142

Improved Weak Classifier Optimization Algorithm

Shaowen Liao and Yong Chen

Abstract In view of the slow speed and time-consuming training problem of the human face detection in complex conditions, we put forward an improved algorithm. To counter the time-consuming training defect of the Adaboot algorithm, we improve the about error rate calculation formula while training the weak classifier, thus accelerating the training speed of the latter and reducing the overall training time. The experimental results show that the improved system has greatly improved the training speed.

Keywords Face detection · The Adaboost algorithm · The about error rate · Weak classifier

1 Introduction

The AdaBoost algorithm is a very important feature classification algorithm in the machine learning, which has been widely used in image retrieval, facial expression recognition and so on [1–3]. The Adaboot algorithm face detection algorithm [4, 5] calculates the Haar-Like features value through using the integral image according to the gray distribution features of the human face and effectively selects the excellent features from a huge number of rectangular features to form the weak classifiers through training; several weak classifiers are superimposed as an

S. Liao (✉)
Information Technology and Media Institute, Hexi University GanSuZhangYe, Hexi, China
e-mail: nini719@163.com

Y. Chen
Artillery Training Base of General Staff, HeBeiXuanHua, HeBei, China
e-mail: 34464822@qq.com

efficient strong classifier by weighted voting; finally each strong classifier cascades for face detection. This series could do more calculation in areas where there may exist a human face while quickly discarding the background region of the image.

2 Adaboost Face Detection Algorithms

The Adaboot algorithm face detection is an algorithm based on the integral image, cascade detector and Adaboost algorithm [6, 7], and the Adaboost algorithm is its core content. The basic idea of this algorithm is to superimpose a huge number of weak classifiers with average classification ability to form a strong classifier. The Adaboost-based face detection system achieves a real-time detection speed addition, which is not only because it takes the integral images to calculate the feature to reduce duplication of computation, but also that it employs the multi-classifier cascading to make discrimination. However, the large number of Haar-Like features makes it take a long training time. Given the above-mentioned factors, we obtain a recurrence formula by transforming the about error rate calculation formula while training the weak classifiers. With this formula, we obtain the about error rate of a sample by making recursive the about error rate of its preceding sample, which could greatly accelerate the training speed of the weak classifiers and effectively improve the speed and accuracy of the face detection.

3 Weak Classifier

Weak classifier is the classifier which makes classification in accordance with some rough judgment. The basic requirement for it is that it can make distinction between the face and non-face image with the about error rate of slightly less than 50 %. Haar-Like rectangle features and weak classifiers correspond to each other. That is, each rectangle feature correspondingly constructs a weak classifier. The weak classifier is described as follows:

S of the same size of the training set of the training samples are known, including P-positive cases, i.e. the face samples; there are N negative examples, i.e. non-face samples. x_i $(i = 1...S)$ Denotes the i-training sample. y_i are used to identify whether it is a human face, when the Sample x_i is a human face, and the corresponding y_i is 1, otherwise y_i is 0. K represents the number of characteristics of the sample. The sample size determines the quantity of the rectangle feature. Therefore, the feature number of the training sample set for each sample is the same, i.e. the K. For each 20 * 20 sample, there are 78,460 features for 5 feature prototypes. A feature of one sample has a fixed feature value. The feature value of the ith sample i.e. x_i is $\{f_1(x_i), f_2(x_i), \ldots, f_j(x_i), \ldots, f_K(x_i)\}$ $(j = 1...K); f_j$ each feature has a two-class classifier. A weak classifier $h_j(x)$ is composed of a feature

f_j, a threshold θ_j and a sign bias p_j (+1 or −1) for indicating the direction of the inequality:

$$h_j(x, f_j, \theta_j, p_j) = \begin{cases} 1 & p_j f_j(x) < p_j \theta_j \\ 0 & other \end{cases} \quad (1)$$

The manifestations of the weak classifiers show that to determine a weak classifier is to determine the threshold value θ, and the biasing direction p determining optimal classifiers involves training and selecting of the weak classifiers. To train a weak classifier is to determine the optimal threshold of this rectangle features in the case of this weight distribution, making this threshold of the weak classifiers become the lowest classification error in all the training samples. Not all of the rectangle features are taken as weak classifiers; instead, the best weak classifiers are selected. To select a best weak classifier is to select the weak classifier whose classification errors of all the training samples rank the lowest in all the weak classifiers. The specific algorithm for determining a weak classifier is shown as follows:

1. Initialization:

   The weak classifiers Error rate: $\varepsilon = $ inf; //Is initialized to infinity
   The total number of features: K
   The total number of samples: S

2. training of the weak classifiers:
   for j = 1: K {// Determine the threshold and error rate for each feature-trained weak classifier and select the feature with the smallest error rate as the optimal weak classifier.

   (a) Sort $(f_j(x))$ // Rank the value of the Feature j of all the samples in the ascending order.
   (b) for i = 1:S {

   i. Determine the error rate of the Sample x_i for the ranked Feature j, i.e. err_i^j according to Formulae (2) (3) (4).

$$\text{Left error rate: } err\_left(i) = \sum_{k=1}^{i-1} w_k(1 - y_k) + \sum_{k=i}^{S} w_k y_k \quad (2)$$

$$\text{Right error rate: } err\_right(i) = \sum_{k=1}^{i-1} w_k y_k + \sum_{k=i}^{S} w_k(1 - y_k) \quad (3)$$

$$err_i^j = \min\{err\_left(i), err\_right(i)\} \quad (4)$$

where w_k is the weight value of the kth sample; when the kth sample is the human face samples, y_k is 1; when the kth sample is the non-human face samples, anonface, y_k is 0.

   ii. Use Formulae (5, 6) to determine the threshold and direction of the error rate err_i^j, i. e.

$$\theta_i^j = f_j(x_i) \tag{5}$$

$$p_i^j = \begin{cases} 1 & err\_left(i) < err\_right(i) \\ -1 & other \end{cases} \tag{6}$$

}

  (c) Determine the error rate of Feature j, i. e. $err^j = \min\{err_i^j\}$ $i = 1\dots S$, and record the corresponding threshold and direction.

}

3. Select the optimal weak classifier
Select the rectangle feature with the smallest err^j error rate the as the optimal weak classifier $\varepsilon = \min\{err^j\}$ $j = 1\dots K$; record the threshold and direction of the weak classifier and the number of the rectangle features.

4 Improved Weak Classifier

The most time-consuming part of the weak classifier training algorithm is to identify the threshold and direction offset of the trained feature, that is, the calculation of Formulae (2, 3). By the formulae, the calculation of the about error rate for each sample needs to be re-accumulated from the very first one, which is an undoubtedly time-consuming task for the training set with thousands of samples. In this paper, we transform Formulae (2, 3) and obtain a recursive formula for calculating the about error rate.

$$err\_left(i-1) = \sum_{k=1}^{i-2} w_k(1-y_k) + \sum_{k=i-1}^{S} w_k y_k$$
$$= w_1(1-y_1) + \dots + w_{i-2}(1-y_{i-2}) + w_{i-1}y_{i-1} + w_i y_i + w_S y_S \tag{7}$$

$$err\_left(i) = \sum_{k=1}^{i-1} w_k(1 - y_k) + \sum_{k=i}^{S} w_k y_k$$
$$= w_1(1 - y_1) + \ldots + w_{i-2}(1 - y_{i-2}) + w_{i-1}(1 - y_{i-1}) + w_i y_i + w_S y_S$$
$$(8)$$

When Formula (8) is minus Formula (7), we have:

$$err\_left(i) = err\_left(i - 1) + w_{i-1}(1 - y_{i-1}) - w_{i-1} y_{i-1}$$
$$= err\_left(i - 1) + w_{i-1} - 2w_{i-1} y_{i-1}$$
$$(9)$$

Similarly, we can get the recursive formula for calculating the right error rate:

$$err\_right(i) = err\_right(i - 1) - w_{i-1} + 2w_{i-1} y_{i-1} \qquad (10)$$

Thus, the transformed formula makes it possible to obtain a recursive the about error rate for a sample from the counterpart of its preceding sample. In the experiment of the present paper, we use Formulae (9, 10) to calculate the about error rate, which improves the training speed of the weak classifier.

5 Experimental Results

This paper uses the the MIT sample database [8, 9] as the basic training base and adds the human face samples with sensitive expressions and obstructions intercepted from the objectionable images to the positive sample base to obtain a training base with 7,188 samples, in which there are 2,807 human face samples and 4,381 non-human face samples.

Before the training, all samples histogram normalization, in order to reduce the image itself, the impact of the gray distribution. Experiment provisions each strong classifier false detection rate $f_{max} = 0.4$, minimum detection rate $d_{min} = 0.99$, the

Table 1 Cascade classifier training results

Layers	Number of negative cases	Number of weak classifiers	Threshold of strong classifiers	Detection rate	False detection rate
1	4381	3	0.5062242	0.9922395	0.3878110
2	1699	10	2.881673	0.9959350	0.3719835
3	632	17	4.901662	0.9940872	0.3829114
4	242	19	5.720353	0.9974132	0.3719008
5	90	20	6.740346	0.9974132	0.3444444
6	31	10	4.254773	0.9970436	0.3548387
7	11	5	3.027746	0.9992609	0.1818182
8	2	1	3.098222	0.9959350	0

cascade classifier target false detection rate $F_{target} = 0.0001$, the 8-layer cascade classifier, as shown in Table 1.

The cascade classifier trained in this paper can produce a better detection effect for front human faces. However, due to the small sample size of the training set and the relatively low proportion of non-face samples, the undetected rate is still relatively high.

6 Conclusions

In this paper, we put forward a weak classifier optimization algorithm. This algorithm, in essence, is to obtain a recurrence/recursive formula by transforming the about error rate calculation formula while training the weak classifiers. With this formula, we obtain the about error rate of a sample by making recursive the about error rate of its preceding sample, which could greatly accelerate the training speed of the weak classifiers. With this formula, we obtain the about error rate of a sample by making recursive the about error rate of its preceding sample, which could greatly accelerate the training speed of the weak classifiers and effectively improve the speed and accuracy of the face detection. The results of the simulated MIT face sample base experiment show that this algorithm can optimize the weak classifiers to improve detection accuracy.

References

1. Viola P, Jones M (2001) Robust Real Time Object Detection. Proceedings of 2nd international work shop on statistical and computational theories of vision
2. Xie H, Fang X, Cai Q (2009) A new weak classifier training method for Adaboost algorithm. J Image Graph 14(11):2411–2415
3. Cheng Sheng TU, Li Li D (2003) The typical algorithm of Adaboost series in boosting family. Comput Sci 30(03):30–35
4. Viola P, Jones M (2001) Rapid object detection using a boosted cascade of simple features. In: Proceedings of IEEE conference on computer vision and pattern recognition, Kauai, Hawai, USA
5. Liang LH, Ai HZ, Xu GY, Zhang B (2002) A survey of human face detection. Chin J Comput 25(05):449–458
6. Freund Y, Schapire R (1999) A short introduction to boosting. J Jpn Soc Artif Intell 14(5):771–780
7. Wei D, Li L (2006) Improvement of Adaboost faces detection. Comput Appl 26(3):619–621
8. Xiong SW, Zong XL, Zhu GF (2005/2006) Improved face detection method based on AdaBoost algorithm. Appl Res Comput
9. Luo MG (2006) The design and implementation of the face detection system based on Adaboost algorithm. Kunming University of Science and Technology

Improvement of the Sustainable Growth Model and Its Application in the Listed Power Companies of China

Li Liu and Zehong Li

Abstract In order to achieve the goal of sustainable growth of finance, the enterprise should always, according to financial sustainable growth rate and its difference from the real growth rates, find out the reason for the difference between them. And then put forward the corresponding management strategy, to realize the sustainable development of the enterprises. Based on the traditional sustainable growth model, this paper made an improvement and reconstruction of it, by constructing the financial sustainable growth model under the condition of the easing of all conditions. And in the paper, we analyzed the reason of the deviation of the present situation of our listed power company's sustainable development from the goal condition, leading to the reasonable suggestions for their sustainable development.

Keywords Listed power company · Sustainable growth · Improvement of the model · Management strategy

1 Introduction

The enterprise's growth and development of are two concepts of relations as well as difference, but with the word "sustainable", their deep meaning will tend to agree [1]. And the financial sustainable growth is one of the most important aspects of the enterprise's sustainable growth, which means the increase of the

L. Liu (✉)
Department of Law, Hebei College of Finance, Baoding, China
e-mail: newtc1989@sina.com

Z. Li
School of Business Administration, North China Electric Power University, Baoding, China
e-mail: 15932139282@163.com

enterprise should coordinate with the enterprise's financial resources situation, that is, on the foundation of the financial policy made by the enterprise itself, taking the enterprise can bear the liability ratio as the premise, get debt capital to support sales growth according to the enterprise's retained earnings growth [2]. Financial sustainable growth is a kind of enterprise growth mode, which uses the endogenous capital and debt capital under the premise of keeping the company debt ability to support enterprise sales growth, but it will not exhaust enterprise wealth [3]. So as to achieve the synchronous growth of company sales and financial resources, and this helps the enterprise's steady development.

2 Traditional Sustainable Growth Models

2.1 Higgins' Financial Sustainable Growth Model and its Shortcomings

Senior US financial expert Robert Higgins professor put forward the concept of enterprise's sustainable growth for the first time, which later was very popular with the companies such as the Boston Consulting Group Company [4]. Professor Henry Higgins mainly made the analysis of sustainable growth from a static view, and put forward the sustainable growth model, of which the basic assumption preconditions are [5]:

(1) Enterprises intend to grow at the growth rate allowed by the market conditions;
(2) Managers are not possible or willing to selling new shares;
(3) Enterprises want to maintain a target capital structure and fixed dividend policy. The model is:

$$
\begin{aligned}
\text{sustainable growth rate} = &\text{ asset turnover} \times \text{net profit margins on sales} \\
&\times \text{equity multiplier} \times \text{retained earnings rate}
\end{aligned}
\tag{1}
$$

The main characteristic of Higgins' sustainable growth model is simple and easy to operate, which briefly generalizes the financial factors influencing the sustainable growth. But the hypothesis of Higgins' sustainable growth model are too harsh, in reality which are difficult to achieve. And this has led to the result that this model, being a static model, cannot reflect the dynamic complex enterprise environment.

2.2 Van Horne's Financial Dynamic Sustainable Growth Model and its Shortcomings

The premise conditions required by Higgins' static model are very demanding, which in real management activities are difficult to meet, so Professor Van Horne put forward that in a given year, the financial sustainable growth is influenced by equity financing and dividend payment factors [6]. Based on this, it is necessary to introduce the sales (S_0) and the shareholders' equity (E_0) of the base period as the foundation of the model which finally turns a static model into a dynamic model:

$$SGR = \frac{\left(E_q + NewE_q - D_{iv}\right)\left(1 + D/E_q\right)(S/A)}{1 - \left[(NP/S)\left(1 + D/E_q\right)(S/A)\right]} \times (1/S_0) - 1 \qquad (2)$$

E_q: Initial shareholder rights; New E_q: Capital newly rose during the last year; D_{iv}: Dividend payable; NP: Net profit; S_0: Previous sales; S: Current sales; D/E_q: Debt/shareholders' equity; A: Total amount of assets.

The advantage of Professor Van Horne's sustainable growth model is that it takes a dynamic perspective, considering the complex enterprise financial environment, which matches the enterprise dynamic business environment. But in the practical application, people would rather continue using Higgins' growth model in the decision. It is because Van Horne's financial sustainable growth model's maneuverability is poorer, and many data are unable to get accurately.

3 Reconstruction of the Financial Sustainable Growth Model

3.1 The Sustainable Growth Model Under the Enterprise's Balanced Growth

Balanced growth means the enterprise keeps the current capital structure and the financial risk related to the structure steady. According to the ratio of the growth of the owner's equity, the enterprise increases its debt, so as to support the sustainable growth of the sales. And at the same time, there is no need to restrict the net profit margin on sales, the asset turnover or the retained earnings rate. On the contrary, under the premise of remaining in the capital structure, according to circumstances and changes in the external environment the enterprise may adjust the net profit margin on sales, the asset turnover and the retained earnings rate.

Under the condition of balanced growth, the enterprise keeps the capital structure remains, so we get:

$$SGR = \frac{A}{(1 - P \times R \times T \times A) \times A_0} - 1 \qquad (3)$$

P: Target net profit margin on sales; R: Target retain yield; T: Target equity multiplier

And compared with Higgins' sustainable growth model, the sustainable growth model under the balance conditions is more flexibility, which is associated with not only the net profit margin on sales, asset turnover, retained earning rate, and rights multiplier but also the of the base period and the target turnover rate. Through the model we can calculate an annual financial sustainable growth rate that is more continued and comparable.

3.2 The Sustainable Growth Model with the Enterprise Capital Structure Changed

The sustainable growth model with the enterprise capital structure changed takes a further relaxes on the restrictions of that the capital structure remains the same in the condition of balanced growth. The model allows enterprise to change its operation efficiency and financial policy with sales growth, and to adjust its capital structure at the same time. Therefore, without secondary offerings, the enterprise can adjust its sales net interest rate, asset turnover, profit retained rate, and capital structure according to the sales growth. And under this condition the sustainable growth model changes as follows:

$$SGR = \frac{A \times T}{(1 - P \times R \times T \times A) \times A_0 \times T_0} - 1 \qquad (4)$$

Compared with the sustainable growth model under the balance conditions the, the model with the enterprise capital structure changed takes a further relax on the restrictions of the steady of the capital structure, and it introduces the equity multiplier of the base period and target equity multiplier, making the model applicable to the enterprise which changes its capital structure due to the sales growth.

4 Application in the China's Listed Power Company

4.1 The Origin of Data and the Basis of Selection

This paper selects the data from the financial statements of 16 listed power companies since 2004. The 16 listed power company are: the Huaneng power international (600011), Datang power (601991), Huadian power international (600027), the Guodian power (600759), Guotou power (600886), Shanghai power (600021), Yuedian Power (000539), Changjiang power (600900), Neimeng power (600863), Sanxia power (600116), Guiguan power (600236), Jingneng power

(600578), Guidong power (600900), Xichang power (000539), Bindian power international (600969), Chuantou power (600674), which is in total of 112 collection of data. Choosing these 16 electric listed companies as the sample is mainly for these 16 electric listed companies' market share is higher and they are distinctly representative, more suitable to reflect the constant change of the structure of the nationwide electricity market.

4.2 Research Method Design of the Financial Sustainable Growth

This paper selects Will Kirk Sen's symbols rank inspection as financial sustainable growth research methods, Will Kirk Sen's symbols rank inspection's goal is mainly to test whether sample company's financial sustainable growth rate and real growth rate is significant.

4.3 The Empirical Results of the Listed Power Companies

The results of Will Kirk Sen's symbols rank inspection test are as follows:

From the Table 1 we can see the sample companies' real growth rate and financial sustainable growth rate in the years from 2004 to 2010, whose Z value's corresponding 2-tailed probability were less than 0.05, so the original hypothesis of no significant differences should be refused to, which means that during 2004–2010 sample companies' real growth rates are not consistent with the financial sustainable growth rates. The sample companies have not realized the sustainable financial growth, and through the Will Kirk Sen's symbols rank inspection test we will get further confirmation about it. This method, through the calculation of sample data absolute deviation's rank and positive and negative value's rank, has considered not only the symbols of the difference of the two groups of sample data, but also the size of the difference. It is a more accurate and

Table 1 Results of Will Kirk Sen's symbols rank inspection test

Year	Z	Asymp. Sig. (2-tailed)
2004	−1.236	0.034
2005	−0.927	0.057
2006	−2.310	0.004
2007	−0.756	0.043
2008	−3.161	0.001
2009	−1.836	0.012
2010	−1.016	0.021
Mixed	−1.606	0.018

Table 2 Results of Will Kirk Sen's symbols rank calculation

		N	Mean rank	Sum of ranks
Sustainable growth rate–real growth rate	Negative ranks	74[a]	56.62	3680.50
	Positive ranks	36[b]	56.33	2647.50
	Ties	0[c]		
	Total	112		

[a] sustainable growth rate < real growth rate
[b] sustainable growth rate > real growth rate
[c] sustainable growth rate = real growth rate

useful method of inspection, which can be used to inspect and confirm the 16 sample companies' further information about their growth situation. The result of its calculation is shown in Table 2.

5 Conclusions

From the results of Will Kirk Sen's symbols rank calculation, we can see that between the selected sample of 16 listed power companies' 2004–2010 financial sustainable growth rate and real growth rate, the difference's positive rank 36, accounting for 32.14 % of the total; While the negative rank number is 74, accounting for 67.86 % of the total; Ties equals zero, namely in the sample data there is no sample whose financial sustainable growth rate and real growth rate equal. Therefore, it can be said that in 2004–2010 years' time the selected 16 listed power companies' financial sustainable growth rates are lower than the actual growth rate, in other words, the sample companies' growth don't match the financial resources which is over speed.

References

1. Liu H (2008) Research of costs incurred in China corporate social responsibility. J Cent Univ Finan Econ (6)
2. Biying Z (2007) Study of the role of the government in solving social cost of industrial enterprisesIII. J Neijiang Technol (17)
3. Cao L (2007) Enterprise's social cost and share. Sci Technol Association BBS (5)
4. Li S (2009) Corporate social responsibility. J Finan Finan (4)
5. Lin R (2006) The rational thinking on corporate social responsibility. J Econ Res Group (8)
6. Jiang ZY (2010) External cost of coal and nuclear power [D]

Nuts Positioning System Based on Machine Vision

Rui Li

Abstract Aiming at the problems of high error ration, high labor intensity and low efficient manual static adjustment of screws and nuts, the system for intelligent nut positioning system on the overload detection part of low-voltage circuit breaker production line was designed. Machine vision technology was taken into the detection unit of the production line. The images of hexagon nut were real-time captured by the CCD camera and processed with image processing technologies such as binarization processing, Gaussian filter, canny edge extraction, Hough transform etc., then the center point position of the Hexagon nut was found, and was used to screws and nuts assembly. Now the system has already been qualified and accepted by factory's customer, and running very well.

Keywords Machine vision · Binarization processing · Gaussian filter · Canny edge extraction · Hough transform

1 Introduction

With the increasing demands of power energy, the required quantity of low-voltage circuit breaker has been developing rapidly during recent years. Manufacturers producing low-voltage circuit breakers have been devoting to improve the efficiency of mass production line, and meanwhile emphasizing on the general parameters control of low-voltage breakers. The efficiency and accuracy of inspection was improved a lot by application of machine vision control, also the automation and intelligent was increased greatly by machine vision control instead

R. Li (✉)
College of Electronic and Electrical Engineering, Shanghai University of Engineering
Science, Shanghai 201620, China
e-mail: reelee111@msn.com
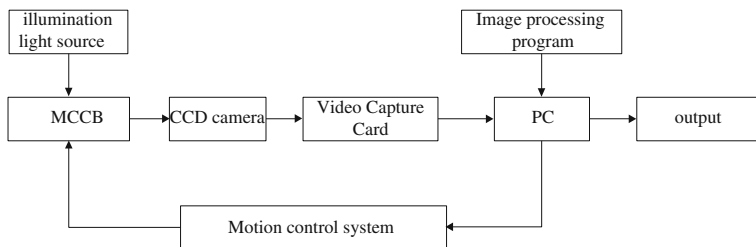
Fig. 1 Fundamental constituent diagram of the nut positioning system

of human checking through eyes [1]. In this system, first the precise location between bolt and nuts in low-voltage breaker for assembly was realized through machine vision inspection technology, then the breaker was tested to trip or non-trip in setting time, so as to the protection of overload during power line [2]. With the help of machine vision inspection system used in bolt and nuts assembly of low-voltage breaker, now the high ration of re-adjustment fault caused by manual assembly, low efficiency and high labor intensity in local manufactures have been solved.

2 Nuts Positioning System

The hardware configuration of nuts positioning test system is composed of test bench, industry personal computer (IPC) and motion control. The test bench includes a LED annular lighting, a CCD camera and a bracket, etc. IPC, the host computer in the mass production line, was installed with video capture software and video capture card. The function of loading and unloading as well as transfer to test station was finished by motion control system through Micro Control Unit [3]. System components chart is shown as Fig. 1:

The modular concept was used in the system's software design, the library Emgu CV which is an encapsulation library of OpenCV's image processing library under .NET platform and used for computer vision processing, was called in programming language C#. The OpenCV, a Intel® Open Source Computer Vision Library [4], consisting of a set of C functions and a few C++ classes, realizing many general algorithms about image processing and computer vision. In this paper, the image pre-processing and feature extraction was implemented by calling the image processing and machine vision algorithms in Emgu CV in C# language. Software modular structure chart is shown a Fig. 2 [5]:
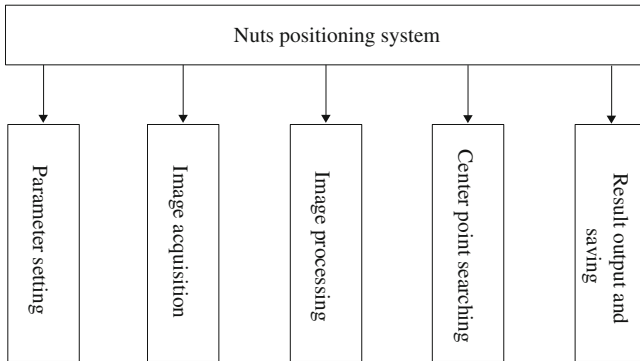
Fig. 2 Software model structure of the visual inspection system

3 Image Acquisition

3.1 Lighting Source Design

The lighting source directly affects the image quality during vision processing procedure. The LED annual lighting was selected in this system, compared with other lighting source, the advantage of annular LED is not only the uniformity and stability of LED brightness, fixed irradiation beam, but also the interference of beam can be reduced by adding polarize and thus improve image quality. Meanwhile, it also equipped with low energy consumption and low heat characteristic, and be suitable for mass production line. The LED lighting source was mounted under the tested nut parts in this system, the difference between hexagonal brass nut and environment can be distinguished obviously, and helpful for next step of image processing.

3.2 CCD Camera

A micro industrial CCD camera was used in this system, and also the dimension and shape of this camera were tailored by reserved space of test bench, The advantages of this camera are low power consumption, long using time, high image quality, good sensitivity and easy to use by USB connection with computer. In this system, the measured optical image of nut was converted into electrical signal, then processed by image capture card and inputted into IPC as digital image, and then was farther processed in IPC [6].

3.3 Image Capture Card

Image Capture Card was connected to the PCI expansion slots in the IPC. In the system, the output analog signal sampled by CCD camera was converted into digital signal by A/D conversion, and then the captured image was directly sent to the host's system memory by high-speed PCI bus. Image capture card is an interface between image acquisition and image processing, regarded as a bridge in machine vision systems.

4 Method of Visual Detection

As the environment influence to shooting, the vision effect and definition of captured image need to be improved through machine vision system. The quality of image processing is mainly depends on whether the featured parameter of tested object can be extracted or not, and whether the image noise linked to the result accuracy can be reduced or not [7]. The inspection process is shown in Fig. 3:

4.1 Gray Treatment and Binarization

In order to discriminate the image of measured object and background environment, the captured colorful image need to be grayed at first, shown in Fig. 4:

Image binarization as a basic image processing technology, can be described intuitively that is to convert a grayed or colorful image into a black and white image. To gray image, through a setting threshold t ($0 \leq t \leq 255$), the image data was divided into two parts: One part which is greater than the threshold was defined as "white", Correspondingly the darker part was "black". Basic steps of binariziation are as follows:

1. Preprocessing gray image, and reduce o eliminate noise by low-pass filtering;
2. Finding the optimal threshold value t by algorithm;
3. The gray value of pixels which are greater than t was defined 255, the lower was 0. Then the target image and background would be detached after the image processed.
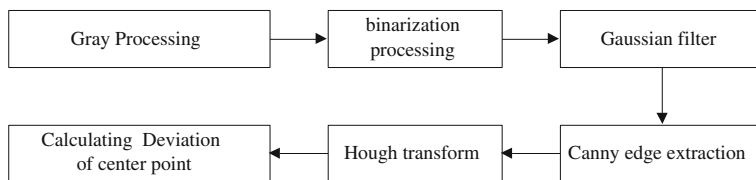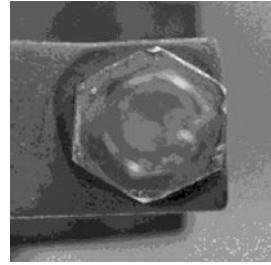

Fig. 3 The flow diagram of the vision inspection

Fig. 4 The picture after gray
processing

4.2 Canny Edge Detection

First with image edge detection, the effects caused by elements as background, etc. can be effectively removed. In this paper, the canny edge detection which is detecting image edge by Canny operator was used. Canny operator is a classical edge detection operator fully reflecting mathematical properties of optimal edge detection, it is an optimal approximation operator for the product between signal noise ratio and positioning, and widely used in image processing and pattern recognition issues [8].

Detailed steps of Canny edge detection are as follows:

1. De-nosing process of image by using the Gaussian filter. In this system, the Gaussian filter and Laplacian sharpening filter was used together to smooth noise firstly, and then detected edges. With the Laplacian filter, the surrounding brightness variation information can be enhanced, the outline of object can be extracted with outstanding details, and therefore, the effect was more desirable [9].
2. The significant change of gray point in adjacent area was displayed extrusive by enhanced algorithm and obtaining the gradient value in the crossbar and the diagonal direction respectively.
3. The point with maximum local gradient was known as the non-maximum value suppression.
4. Implementing edge detection and link by setting the threshold of low gradient and high gradient, the result of the processing was shown in Fig. 5:

Fig. 5 The picture after
canny edge extraction

4.3 Hough Algorithm

The edge of hexagonal nut has been weakened nearly into a circle after above operations. So the center searching of the hexagonal nut can be transformed into looking for the center of the circle. Now Hough transformation is the most widely used method to circle detection. The essential of Hough transformation is accumulating the pixels which have certain relationships in image space, and searching corresponding points which can link pixels in parameter space with an analytic form [10]. To some discontinuous conditions caused by noise interference and covered by other target in area edge, as the complete described parameters were calculated by local measurement, it has good fault tolerance and robustness. However, once the parameters space increased, the calculation quantity will rise quickly, simultaneous the storage space will be consumed largely, and time used also will surge. The OpenCV Hough improved algorithm was used in this system, the algorithm is not only retaining the advantage that the standard Hough transform is not sensitive to noise, but also increasing speed by using the circularly symmetric geometric features to detect the center [11]. Two steps of detail detection method:

1. Since the center located in gradient direction of each point in contour line, the center can be determined by using gradient information through Hough transformation's 2D voting and local peak value.
2. Determining the radius through one-dimensional transform.

   Algorithm steps to calculate the center abscissa are as follows [12]:

1. finding satisfied the following conditions p_{xj} by traversal method: the pixels are not edge points; searching leftward and rightward from the pixels p_{xj} adjacent left and right point, until find the first left and right adjacent edge pixel, referred to as r_{xp} (right searching) and l_{xq} (left searching); $p_{xj} - l_{xq} = r_{xp} - p_{xj}$.
2. Count accumulating to the above satisfied conditions p_{xj} by Hough transform in the one-dimensional space.
3. The p_{xj} corresponding to the maximum accumulated count in parameters space is the abscissa of the circle center: $a = p_{xj}$

   Similarly, the vertical coordinate (b) of the center can be calculated $b = p_{ix}$, thus the center of the circle is: (a, b)

   The arithmetic of the radius is using the obtained center coordinates (a, b), substitute the edge pixels p_{xj} into the equation of a circle $(x - a)^2 + (x - b)^2 = r^2$, and calculate the candidate radius r, then accumulate count to candidate radius by Hough transform in the one-dimensional space. Finally, determine the circle by judging whether the count value of r is more than the minimum number of points can be constituted a circle. The processed image is shown in Fig. 6:

Fig. 6 The picture after
Hough transforms
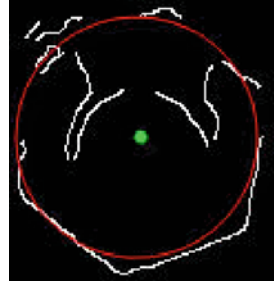
4.4 Calculating the Offset Value

The center point of the hexagon nut (a, b) was obtained after above operation, but there an offset value d_x, d_y compared with the obtained reference point (x, y) obtained:

$$d_x = a - x; \tag{1}$$

$$d_y = b - y; \tag{2}$$

Then calculating the actual moving distance of test platform controlled by PLC D_x, D_y

$$D_x = k * d_x + m; \tag{3}$$

$$D_y = k * d_y; \tag{4}$$

Wherein, k is the ratio value between object and image; m is the relative distance between optical center of the camera and the center of the bolt in the x-axis.

5 System Implementation

Image capture was finished by the host computer in the long delay unit of circuit breaker mass production line the video capture card was installed in the PIC slot of IPC. The communication between IPC and PLC was realized by utilizing OPC technology through C# programming. The vision detecting was started in IPC after image capture, and then the processed center point and coordinate information were sent to PLC, final the assembly position between bolt and nuts were adjusted by PLC according the received real-time parameter, till finishing assembly. The interface of host computer based on C# is shown in Fig. 7:

Fig. 7 The system interface based on visual C#

6 Conclusions

In this paper, the machine vision technology was adopted to achieve intelligent assembly of bolt and nut in the general parameters detection unit during low-voltage circuit breaker mass production line, and thus fulfilled overload protection test of circuit breaker [13]. Through analysis of technical rules and actual environment in mass production line, and continuously updating and debugging of the machine vision system, the assembly of bolt and nut can be completed, and also the real-time and accurate overload protection test can be passed during trial-run period in factory [14]. The system can meet actual operational needs during mass production, and is an important automation progress for intelligent measure system of low-circuit breaker.

References

1. Steger C, Ulrich M (2008) Algorithm and application in machine vision. Yang S. Tsinghua Publishing House, Beijing, pp 164–350
2. Dawoud A, Kamel M (2002) Iterative model-based binarization algorithm for cheque images. Int J Doc Anal Recogn. doi:10.1007/s10032-002-0081-9
3. Zhu Q, Liu H, Chen B (2012) An improved method for multiple targets tracking. J Cent South Univ. doi:10.1007/s11771-012-1351-4
4. Bradski G, Kaehler A (2009) Study OpenCV (chinese version). Tsinghua Publishing House, Beijing, pp 173–182
5. Wu J, Bin H (2010) Research and development of dimensional inspection system for thin sheet part based on machine vision. Mach Tool Hydraulics 09:86–88
6. Kazanskiy NL (2010) Machine vision system for singularity detection in monitoring the long process. Opt Mem Neural Networks 19(1):22–30

7. Gu Y, Mingxin He (2011) Research on beer bottle detection system based on machine vision. Comput Eng Appl 01:248–252
8. Wang T, Xu Y, Kang H (2010) Image binarization based on canny operator. Microcomputer applications 26:4–7
9. Ukwatta E, Samarabandu J, Hall M (2011) Machine vision system for automated spectroscopy. Mach Vis Appl A. doi:10.1007/s00138-011-0338-8
10. Xiao Z, Wei P (2006) Detection of circle based on Hough transform. Transducer Microsyst Technol 62–64
11. Roshchin D (2012) Increasing the efficiency of systems for the detection of the contours of parts on the basis of Hough's transformation. Meas Tech. doi:10.1007/s11018-012-9896-z
12. Shang L, Li R, Song X (2011) Improved Hough transform circle detection algorithm. Electron Des Eng 07:86–88
13. Li X, Zhang Y, Wang W (2011) System design of processing error detection based on machine vision. Comput Eng Appl 48(4):71–73
14. Huang L, Zhang Y, Hu B (2012) Automatic detection of liquid level in transparent bottle based on machine vision. Autom Instrum 2:57–60

Information Fusion Estimation Based Optimal Control for Ship Course Control Problem

Guoqing Xia and Huiyong Wu

Abstract In this paper, a ship model with wave disturbances is used to simulate the ship autopilot system. An information fusion estimation based optimal tracking controller is designed for the ship course tracking system. Due to the effect of wave disturbances, the passive observer is adopted to estimate the system state and extern disturbance. Finally, two simulation tests in calm water and in waves respectively are carried out to demonstrate the effectiveness of the proposed control scheme.

Keywords Information fusion control · Autopilot · Passive observer

1 Introduction

Information fusion estimation mainly uses the concept of optimal estimation, which commonly exists in all kinds of decision problems. It can be considered as a decision problem when the control variable is in the type of discrete. When the control variable is considered as the decision, the optimal solution of the control variable can be obtained by using the optimal estimation method [1]. Based on information fusion estimation, the author of [1] discussed optimal control with expected trajectory, optimal preview control, optimal tracking control and LQ optimal control in [2–5]. In these papers, however, the system states or the extern disturbances are all known, while in practice, these are not always satisfied.

For the navigation of ships, the autopilot control system has been widely used to make the ship sail in the commanded path [6–13]. Fossen discussed the nonlinear weather optimal control of ships exposed to environmental disturbances in [6].

G. Xia (✉) · H. Wu
College of Automation, Harbin Engineering University, Harbin, China
e-mail: wuhyong1688@163.com

In [7], a predictive controller is designed to steer the ship sailing forward with the constant velocity along the predefined reference path. In [8], Ming-Chung Fang adopted sliding mode control technique and line-of-sight guidance technique to navigate the ship. Fossen and Perez designed the DP controller by using Kalman filter and the conventional PID control method in [9], to name a few.

In this note, the concept of information fusion control is applied on the ship steering problem. To compensate the wave disturbances, a passive observer is adopted to estimate the system state and extern disturbance. The stability of controller-observer system can be proven by applying the separation principle.

The structure of the paper is as follows. In Sect. 2, the ship course optimal tracking task is formulated. Section 3 designs the passive observer. Section 4 contains the design of information fusion based optimal tracking controller. Section 5 provides simulation results for the proposed control scheme. Finally, conclusions are made in the last section.

2 Problem Formulation

Assume that the ship has an xz-plane of symmetry; heave, pitch roll modes are neglected; the body-fixed frame co-ordinate origin is set in the center-line of the ship; the speed in surge is constant; the ship is controlled by a single rudder; the rolling mode is negligible. Then the ship model in [13] can be reduced to the following autopilot model:

$$\dot{\boldsymbol{\eta}} = \mathbf{R}(\boldsymbol{\eta})\mathbf{v} \tag{1}$$

$$\mathbf{M}_1\,\dot{\mathbf{v}}_1 = -\mathbf{N}(u_0)\mathbf{v}_1 + \mathbf{b}_1 u + \mathbf{E}_1(\psi)\boldsymbol{\varepsilon} \tag{2}$$

With $\boldsymbol{\eta} = [x, y, \psi]^{\mathrm{T}}$, $\mathbf{v}_1 = [v \quad r]^{\mathrm{T}}$, $\mathbf{b}_1 = [-Y_\delta \quad -N_\delta]^{\mathrm{T}}$,

$$\mathbf{M} = \begin{bmatrix} m_{11} & 0 & 0 \\ 0 & m_{22} & m_{23} \\ 0 & m_{32} & m_{33} \end{bmatrix}, \mathbf{C}(\mathbf{v}) = \begin{bmatrix} 0 & 0 & -m_{22}v - m_{23}r \\ 0 & 0 & m_{11}u \\ m_{22}v + m_{23}r & -m_{11}u & 0 \end{bmatrix},$$

$$\mathbf{R}(\boldsymbol{\eta}) = \mathbf{R}(\psi) = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{D} = \begin{bmatrix} -X_u & 0 & 0 \\ 0 & -Y_v & -Y_r \\ 0 & -N_v & -N_r \end{bmatrix}.$$

$$m_{11} = m - X_{\dot{u}}, \ m_{22} = m - Y_{\dot{v}}, \ m_{33} = I_z - N_{\dot{r}}, \ m_{23} = m_{32} = mx_g - Y_{\dot{r}}.$$

Where (x, y, ψ) denote the coordinates and heading of the ship in the earth-fixed frame; v and r denote the speeds in sway and yaw in the body-fixed frame; $u = \delta$ is the control input; $\mathbf{M}_1$, $\mathbf{N}(u_0)$, and $\mathbf{E}_1(\psi)$ are derived from $\mathbf{M}$, $\mathbf{C}(\mathbf{v})$, $\mathbf{R}(\psi)$ and $\mathbf{D}$; $\mathbf{C}(\mathbf{v})$ is the Coriolis-centripetal matrix, $\mathbf{D}$ is the damping matrix, $\mathbf{R}(\psi)$ is a state dependent transformation matrix; ε is a bias term representing slowly varying

environmental forces and moments; m is the mass of the ship; I_z is the ship's inertia about the Z_b-axis of the body-fixed frame.

Combining $\dot{\psi} = r$ and (2), yields:

$$\dot{\mathbf{x}} = \mathbf{A}_c \mathbf{x} + \mathbf{b}_c u + \mathbf{E}_c(\psi)\boldsymbol{\varepsilon} \tag{3}$$

$$\mathbf{y} = \psi = \mathbf{C}\mathbf{x} \tag{4}$$

where $\mathbf{x} = [\,v \quad r \quad \psi\,]^{\mathrm{T}}$, $\mathbf{A}_c$, $\mathbf{b}_c$ and $\mathbf{E}_c(\psi)$ can be calculated by $\mathbf{M}_1$, $\mathbf{N}(u_0)$, $\mathbf{b}_1$, and $\mathbf{E}_1(\psi)$, and $\mathbf{C} = [\,0 \quad 0 \quad 1\,]$.

Control objective: Design the control input u to guide the ship (3) to sail to command waypoints while minimizing the performance index:

$$J = \min_{\mathbf{u}}\left\{\mathbf{e}^T(T)Q_f\mathbf{e}(T) + \int_0^T (\mathbf{e}^T Q\mathbf{e} + \mathbf{u}^T R\mathbf{u})dt\right\} \tag{5}$$

where $\mathbf{e} = \mathbf{y} - \mathbf{y}_d = \psi - \psi_d$, ψ_d is the desired heading which can be obtained for all time $t \in [0, T]$, $Q = Q^T \geq 0$ and $R = R^T > 0$ are the tracking error and control weighting matrices, respectively. The weight matrix $Q_f = Q_f^T \geq 0$ can be included to add penalty on the final state.

3 Passive Observer Design

The conventional filtering of wave disturbances is to apply a Kalman filter [9]. However, the main drawback of Kalman filter is that a relatively large number of parameters must be determined through experimental testing of the ship. In this section, the method of passive observer [14] is used to estimate the unknown variables, and some modification is made to adapt to the presented ship model.

Refer to the observer form in [14], the observer equations can be chosen as:

$$\dot{\hat{\boldsymbol{\xi}}} = \mathbf{A}_w\hat{\boldsymbol{\xi}} + \mathbf{K}_1\tilde{\mathbf{z}} \tag{6}$$

$$\dot{\hat{\boldsymbol{\eta}}} = \mathbf{R}(\mathbf{z})\hat{\mathbf{v}} + \mathbf{K}_2\tilde{\mathbf{z}} \tag{7}$$

$$\dot{\hat{\boldsymbol{\varepsilon}}} = -T^{-1}\hat{\boldsymbol{\varepsilon}} + \mathbf{K}_3\tilde{\mathbf{z}} \tag{8}$$

$$\mathbf{M}_1\dot{\hat{\mathbf{v}}} = -\mathbf{N}(u_0)\hat{\mathbf{v}} + \mathbf{b}_1 u + \mathbf{E}_1(\psi)\hat{\boldsymbol{\varepsilon}} + \mathbf{E}_1(\psi)\mathbf{K}_4\tilde{\mathbf{z}} \tag{9}$$

$$\hat{\mathbf{z}} = \hat{\boldsymbol{\eta}} + \hat{\boldsymbol{\eta}}_w = \hat{\boldsymbol{\eta}} + \mathbf{C}_w\hat{\boldsymbol{\xi}} \tag{10}$$

where $\boldsymbol{\varepsilon} \in \Re^3$ denotes the bias forces and moments; $\boldsymbol{\xi} \in \Re^6$, and $\boldsymbol{\eta}_w$ is the ship's WF motion due to 1st-order wave-induced disturbances; $\tilde{\mathbf{z}} = \mathbf{z} - \hat{\mathbf{z}}$ is the estimation error, $\mathbf{K}_1$, $\mathbf{K}_2$, $\mathbf{K}_3$, and $\mathbf{K}_4$ are the observer gain matrices.

Let $\boldsymbol{\eta}_0 = \begin{bmatrix} \boldsymbol{\xi}^T & \boldsymbol{\eta}^T \end{bmatrix}^T$, and define the estimation errors as $\tilde{\mathbf{v}} = \mathbf{v} - \hat{\mathbf{v}}, \tilde{\boldsymbol{\varepsilon}} = \boldsymbol{\varepsilon} - \hat{\boldsymbol{\varepsilon}}$ and $\tilde{\boldsymbol{\eta}}_0 = \boldsymbol{\eta}_0 - \hat{\boldsymbol{\eta}}_0$, then the error dynamics can be written as:

$$\dot{\tilde{\boldsymbol{\eta}}}_0 = [\mathbf{A}_0 - \mathbf{K}_0 \mathbf{C}_0] \tilde{\boldsymbol{\eta}}_0 + \mathbf{B}_0 \mathbf{R}(\boldsymbol{\eta}) \tilde{\mathbf{v}} \tag{11}$$

$$\dot{\tilde{\boldsymbol{\varepsilon}}} = -T^{-1}\tilde{\boldsymbol{\varepsilon}} - \mathbf{K}_3 \tilde{\mathbf{z}} \tag{12}$$

$$\mathbf{M}_1 \dot{\tilde{\mathbf{v}}} = -\mathbf{N}(u_0)\tilde{\mathbf{v}} + \mathbf{E}_1(\psi)\tilde{\boldsymbol{\varepsilon}} - \mathbf{E}_1(\psi)\mathbf{K}_4 \tilde{\mathbf{z}} \tag{13}$$

where $\mathbf{A}_0 = \begin{bmatrix} \mathbf{A}_w & 0 \\ 0 & 0 \end{bmatrix}$, $\mathbf{B}_0 = \begin{bmatrix} 0 \\ \mathbf{I} \end{bmatrix}$, $\mathbf{C}_0 = [\, \mathbf{C}_w \quad \mathbf{I} \,]$, and $\mathbf{K}_0 = \begin{bmatrix} \mathbf{K}_1 \\ \mathbf{K}_2 \end{bmatrix}$.

The passivity and stability of the observer will be provided if the observer gain matrices $\mathbf{K}_1, \ldots, \mathbf{K}_4$ satisfy the KYP lemma [14], and the observer gain matrices should be chosen to have a diagonal structure.

4 Optimal Control for the Ship Course Tracking

In this section, an optimal tracking controller is designed. The heading autopilot control systems are incorporated with Eqs. (3–4) to simulate the ship course tracking problem. According to the LOS guidance [8], the desired heading angle is:

$$\psi_d = \mathrm{atan2}(y_{wp} - \hat{y}, x_{wp} - \hat{x}) \tag{14}$$

Here, $(\hat{x}, \hat{y})$ is the estimated coordinate of the ship whereas (x_{wp}, y_{wp}) is the coordinate of the waypoint position. Each waypoint has its acceptable radius which is in general around one to three ship lengths. When the ship moves into the acceptable radius, the ship will be required to sail to the next waypoint.

As mentioned before, the information fusion estimation technique can be introduced into the optimal control when the control variable is in the type of discrete. Therefore, firstly we should get the discrete form of the model (3–4), that is:

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{b}u(k) + \mathbf{E}\boldsymbol{\varepsilon}(k) \tag{15}$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) \tag{16}$$

where $\mathbf{A}, \mathbf{b}$ and $\mathbf{E}$ can be calculated through the discretization.

Meanwhile, the performance index (5) is changed to the discrete form:

$$J = \min_{\mathbf{u}} \left\{ \sum_{k=0}^{N-1} \mathbf{e}^T(k+1)Q(k+1)\mathbf{e}(k+1) + \sum_{k=0}^{N-1} \mathbf{u}(k)^T R(k)\mathbf{u}(k) \right\} \tag{17}$$

where $Q(k)$ and $R(k)$ denote information weight [1] of the tracking error and control quantity respectively.

Notice that the system state $\mathbf{x}(k)$ and the disturbance term $\boldsymbol{\varepsilon}(k)$ in Eqs. (15–16) should use the estimated values derived from the observer, and the denotation signs of them are not changed to avoid confusion with the following parts.

Here we use the iterative algorithm of optimal estimation to calculate the optimal control strategy, since the iterative method is equivalent to close-loop control and has less computational complexity [2, 4].

Firstly, suppose that at the kth moment all the information has been fused; optimal estimation $\hat{\mathbf{x}}(k+1)$ of the co-state and its information weight $P(k+1)$ have been obtained. Then we can get the measurement equations of the control variable from the system state equation and the performance index [4]:

$$0 = u_d(k) = u(k) + n(k) \tag{18}$$

$$\hat{\mathbf{x}}(k+1) = \mathbf{x}(k+1) + \mathbf{w}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{b}u(k) + \mathrm{E}\boldsymbol{\varepsilon}(k) + \mathrm{w}(k+1) \tag{19}$$

where $n(k)$ and $\mathbf{w}(k+1)$ are all zero-mean white noises, with variation of $R^{-1}(k+1)$ and $P^{-1}(k+1)$ respectively.

Then, according to the information fusion estimation theory [1], we can obtain the optimal solution of the control strategy by fusing its two measurement equations mentioned above, that is:

$$\hat{u}(k) = \left[R(k) + \mathbf{b}^T P(k+1)\mathbf{b}\right]^{-1} \left\{\mathbf{b}^T P(k+1)[\hat{\mathbf{x}}(k+1) - \mathbf{A}\mathbf{x}(k) - \mathrm{E}\boldsymbol{\varepsilon}(k)]\right\} \tag{20}$$

Secondly, the acquisition of the co-state $\hat{\mathbf{x}}(k+1)$ and its information weight $P(k+1)$ is discussed. Similarly, suppose that all the state estimation and information weight after the $(k+1)$th moment have been fused, then we can obtain the measurement equations of the co-state as follows [2]:

$$\hat{\mathbf{x}}(k+2) = \mathbf{A}\mathbf{x}(k+1) - \mathbf{b}n(k+1) + \mathrm{E}\boldsymbol{\varepsilon}(k+1) + \mathrm{w}(k+2) \tag{21}$$

$$\mathbf{y}_d(k+1) = \mathbf{C}\mathbf{x}(k+1) + m(k+1) \tag{22}$$

where $m(k+1)$ is a zero-mean white noise, with variation of $Q^{-1}(k+1)$.

By fusing these two measurement equations, the estimation and information weight of $\mathbf{x}(k+1)$ can be derived as:

$$\begin{aligned}
\hat{\mathbf{x}}(k+1) &= P^{-1}(k+1)\{\mathbf{C}^T Q(k+1)\mathbf{y}_d(k+1) \\
&\quad + \mathbf{A}^T [P^{-1}(k+2) + \mathbf{b}R^{-1}(k+1)\mathbf{b}^T]^{-1} \\
&\quad \times [\hat{\mathbf{x}}(k+2) - \mathrm{E}\boldsymbol{\varepsilon}(k+1)]\}
\end{aligned} \tag{23}$$

$$P(k+1) = \mathbf{C}^T Q(k+1)\mathbf{C} + \mathbf{A}^T \left[P^{-1}(k+2) + \mathbf{b}R^{-1}(k+1)\mathbf{b}^T\right]^{-1} \mathbf{A} \tag{24}$$

Equations (23) and (24) are solved in the inverse-time direction. During the solution, it is necessary to know all the future ideal outputs and disturbances. However, in the case of ship tracking through the commanded waypoints in waves, only the current information can be obtained, and both the future ideal outputs and

disturbances are unknown. The author of [4] illustrated how to solve this problem. And formula (24) can be changed to the following form:

$$P = \mathbf{C}^T Q \mathbf{C} + \mathbf{A}^T \left[P^{-1} + \mathbf{b} R^{-1} \mathbf{b}^T \right]^{-1} \mathbf{A} \tag{25}$$

Then the approximately optimal control law is derived as:

$$
\begin{aligned}
\hat{u}(k) = & - (R + \mathbf{b}^T P \mathbf{b})^{-1} \mathbf{b}^T P \mathbf{A} \cdot \mathbf{x}(k) \\
& + (R + \mathbf{b}^T P \mathbf{b})^{-1} \mathbf{b}^T P \left[P - \mathbf{A}^T (P^{-1} + \mathbf{b} R^{-1} \mathbf{b}^T)^{-1} \right]^{-1} \times \mathbf{C}^T Q \cdot \mathbf{y}_d(k) \\
& - (R + \mathbf{b}^T P \mathbf{b})^{-1} \mathbf{b}^T P \left\{ \left[P - \mathbf{A}^T (P^{-1} + \mathbf{b} R^{-1} \mathbf{b}^T)^{-1} \right]^{-1} \right. \\
& \left. \times \mathbf{A}^T (P^{-1} + \mathbf{b} R^{-1} \mathbf{b}^T)^{-1} \mathbf{E} + \mathbf{E} \right\} \cdot \boldsymbol{\varepsilon}(k)
\end{aligned}
\tag{26}
$$

where $\mathbf{x}(k)$ is the current state, $y_d(k)$ is the current desired output, $\boldsymbol{\varepsilon}(k)$ is the current disturbance, and P is positive definite solution of Eq. (25).

The Eq. (26) indicates that the proposed approximately optimal control contains feedback control of the system state, feed-forward control of the desired outputs, and feed-forward control of the disturbances. Figure 1 shows the structure of the ship course tracking control system, where the observer is described in Sect. 2.

5 Simulation Results

In this section, simulation results will be presented to verify the effectiveness of the proposed control scheme. The results of two simulation tests are given, which include the cases of tracking in calm water and in waves.

In the first case, the ship tracks the desired path given with four waypoints in calm water, and simulation results are shown in left parts of Figs. 2, 3. From the figures, it can be concluded that the optimal autopilot controller can successfully guide the ship through the given waypoints, while the time response of rudder angle is mostly smooth but with some sharp lines, which is due to the maximum rudder and rate limit. Figure 3 depicts the time responses of swaying velocity, angular velocity in yawing and heading angle, which are also smooth and satisfied.

The second simulation test demonstrates the behavior of proposed controller when wave disturbances act on the ship. The simulation results in waves are shown in right parts of Figs. 2, 3. In this case, the key parameters of designed controller are kept unchanged, and the 1-order waves are filtered by the passive observer, in
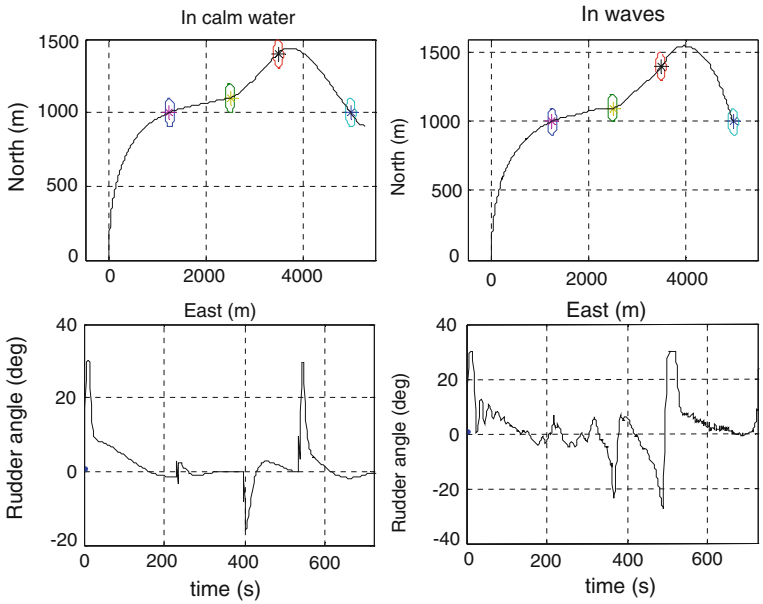

Fig. 1 The ship course tracking control systems
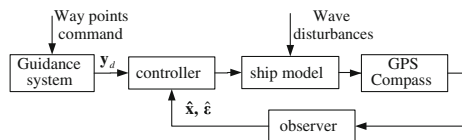
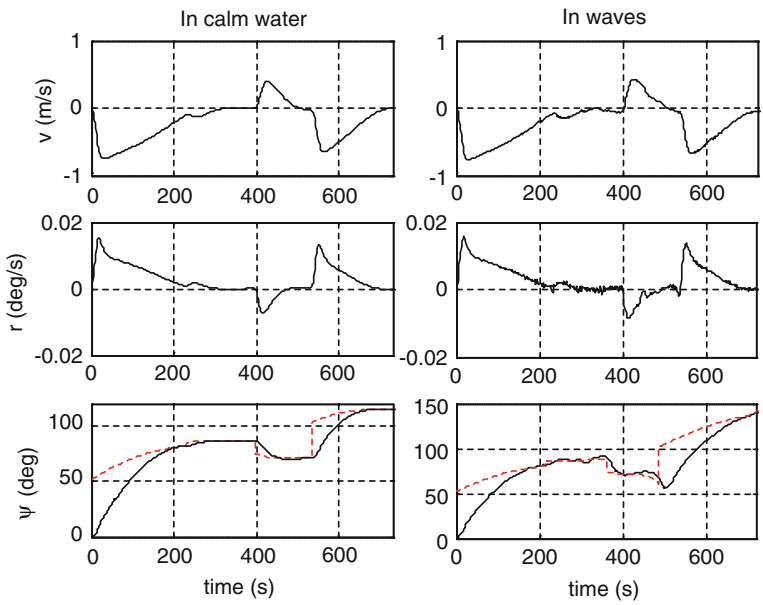**Fig. 2** Waypoint guidance and time response of rudder angle in two cases

Fig. 3 The time responses of the system states in two cases

this way, the influence of 2-order shift wave forces is rejected by the controller. Figure 2 shows that the ship is pushed away along the wave direction (back right 30° in the earth-fixed frame) but still passes through the allowance region, and the time response of rudder angle is also acceptable despite of some small amplitude oscillations. In Fig. 3, the responses of swaying velocity, angular velocity in yawing and heading angle also oscillate with small amplitude, but that is acceptable in engineering field. On the other hand, from the performance index of the two cases, which are 49775.35 and 81254.24, respectively, we can also see the difference of the two tests, and higher the performance index is, more control input is needed and more difficult it is to operate the ship.

6 Conclusion

The information fusion estimation technique is introduced to the ship control system, and an optimal tracking controller based on the information fusion estimation method is designed for the ship course tracking problem. The unknown system states and disturbances are estimated by a passivity observer, and the separation principle guarantees the stability of the observer-controller system. The information fusion control also can be applied on the nonlinear systems, which is more similar to the actual system. With the development of information fusion technique, it is believed that information fusion control is gaining more and more progress.

References

1. Wang ZS (2004) Information fusion control theory and method. Post-doctoral research work report, Nanjing University of Aeronautics and Astronautics
2. Wang ZS, Wang DB (2006) Optimal control with ideal control strategy and expected trajectory. Control Decis 21(1):100–103
3. Zhen ZY, Wang ZS, Wang DB (2010) Information fusion estimation based preview control for discrete linear system. Acta Automatica Sinia 36(2):347–352
4. Zhen ZY, Wang ZS, Wang DB (2009) Discrete system optimal tracking control based on information fusion estimation. Control Decis 24(1):81–85
5. Zhen ZY, Jiang Ju, Wang ZS et al (2012) Finite-time information fusion for linear quadratic optimal control. Control Theory Appl 29(2):172–176
6. Fossen TI, Strand JP (2001) Nonlinear passive weather optimal positioning control (WOPC) system for ships and rigs: experimental results. Automatica 37:701–715
7. Wu J, Peng H, Ohtsu K et al (2012) Ship's tracking control based on nonlinear time series model. Appl Ocean Res 36:1–11
8. Fang MC, Luo JH (2005) The nonlinear hydrodynamic model for simulating a ship steering in waves with autopilot system. Ocean Eng 32:1486–1502
9. Fossen TI, Perez T (2009) Kalman filtering for positioning and heading control of ships and offshore rigs. IEEE Control Syst Mag 32:32–44

10. Gao C, Gao J (2005) On the existence of an optimal control of ship automatic steering instruments. J Ocean Univ China 4(2):185–188
11. Liao Y, Wan L, Zhuang J (2011) Backstepping dynamical sliding mode control method for the path following of the underactuated surface vessel. Procedia Eng 15:256–263
12. Zhang LJ, Jia HM, Qi X (2011) NNFFC-adaptive output feedback trajectory tracking control for a surface ship at high speed. Ocean Eng 38:1430–1438
13. Velagic J, Vukic Z, Omerdic E (2003) Adaptive fuzzy ship autopilot for track-keeping. Control Eng Pract 11:433–443
14. Fossen TI, Strand JP (1999) Passive nonlinear observer design for ships using Lyapunov methods: Experimental results with a supply vessel. Automatica AUT-35(1):3–16

Image Feature Extraction Based on Support Vector Machine and Multi-DSP Combined Structure

Hui Chen and Jiao Hu

Abstract An image feature extraction method based on support vector machine (SVM) is presented in this paper, which first seeks the optimal separating hyperplane in small samples and then projects image data in the corresponding normal direction. In multiclass cases, the method has an optimal choose for selecting projecting axis by some sub-SVMs with simplified structure. A multi-DSP combined structure system has been designed to implement this method by TMS320DM648 and TMS320DM6446. The results show the proposed method is effective, and also meets the real-time requirement.

Keywords Support vector machine · Feature extraction · Multi-DSP · Combined structure

1 Introduction

In image recognition, it is not appropriate for an image which is processed directly due to high dimension original data, especially for real-time images. Therefore, the feature extraction is a very important part in an image recognition system. At present, there are some feature extraction methods used widely in image recognition, such as independent component analysis (ICA) [1], principal component analysis (PCA), Fisher linear discriminate analysis (FDA) [2] and so on. These methods are mostly linear and there are many disadvantages for the typical nonlinear image data. In order to solve this problem, some scholars introduced nonlinear transformations, which convert original image data into a high-dimensional space through nonlinear kernel mapping, and then use PCA, FLD for feature extraction, namely kernel

H. Chen (✉) · J. Hu
School of Computer Science and Engineering, Anhui University of Science and Technology, Huainan 232001, China
e-mail: huichen@aust.edu.cn

principal component analysis (KPCA) [3] and kernel Fisher discriminate analysis (KFD) [4]. These methods utilize samples to estimate mean and variance instead of normal distribution parameters based on the assumption of the original data being normal distribution. If these parameters estimation is not accurate enough, the method would be failure. Therefore, these methods cannot be popularized widely. In feature extraction by utilizing nonlinear methods directly, some representative methods are local linear embedding (LLE) [5] and isometric feature mapping (ISOMAP), especially the LLE method not only has the advantages of translation, rotation invariance etc., but also can keep the topological structure of original data. However, they would introduce a lot of noise in feature extraction and take some adverse impact on subsequent recognition.

The effective feature extraction method should be able to make an original complex problem simplified, and also improve the efficiency of solving problems, which is remarkably important for a real-time recognition system. The image feature extraction method based on support vector machine is presented in this paper, which utilizes many inner excellent properties in support vector machine algorithm, such as good outreach capacity to unknown samples. In addition, a multi-DSP combined structure system has been designed to realize this method in real time.

2 Support Vector Machine

Support vector machine is a learning method based on the structural risk minimization principle [6]. The principle of support vector machine algorithm is expressed briefly as follows: Looking for a optimal separating hyperplane under the conditions of limited samples, namely seeking the optimal discriminate function to make the classification interval be maximum; for linear inseparable problems, especially for high-dimensional nonlinear image recognition problems, data in low dimension space are mapped into a high dimension space by using a kernel function, so as to transform a low-dimensional inseparable problem into a high-dimensional linear separable problem.

Assume that $\{x_i \ y_i\}$ $(i = 1, 2, \ldots, n)$ are training sample data, where $x_i \in R^d$ is d dimensional training sample, $y_i \in \{-1, 1\}$ is classification information of i sample. Support vector machine is to seek an optimal separating hyperplane: $x^T \bullet w + b = 0$. In order to make samples to be separated as much as possible, classification interval should be maximum, that is, minimize the objective function $\|w\|^2/2$. Then, the problem of seeking optimal separating hyperplane is transformed into the following optimization problem:

$$\min \frac{1}{2} \|w\|^2 \tag{1}$$

$$s.t. \, y_i(x^T \bullet w + b) - 1 = 0, \quad i = 1, 2, \ldots, n \tag{2}$$

In order to reduce influence of isolated points on classification and improve training and recognizing speed, Eq. (1) is usually transformed into corresponding dual problem:

$$\max T_D = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j}^{n} \alpha_i \alpha_j y_i y_j x_i x_j \tag{3}$$

$$s.t. \sum_{i=1}^{n} a_i y_i = 0 \tag{4}$$

$$0 \le \alpha_i \le C, \quad i = 1, 2, \ldots, n \tag{5}$$

Since image data is typically multiclass nonlinear, for face expression recognition, image data need to be mapped into a high-dimensional feature space by nonlinear mapping $\Phi : x_i \rightarrow \Phi(x_i)$. Gaussian radial basis function $k(x_i, y_j)$ is selected to realize the inner product of two vector $\Phi(x_i)$, $\Phi(x_j)$, which can be written as:

$$k(x_i, y_j) = \Phi(x_i) \bullet \Phi(x_j) = e^{-\frac{\|x_i - x_j\|}{2\sigma^2}} \tag{6}$$

Hence, the optimization problem of Eq. (3) can be converted into

$$\max T_D = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j}^{n} \alpha_i \alpha_j y_i y_j k(x_i, y_j) \tag{7}$$

$$s.t. \sum_{i=1}^{n} a_i y_i = 0 \tag{8}$$

$$0 \le \alpha_i \le C, \quad i = 1, 2, \ldots, n \tag{9}$$

Then, projection axis w can be obtained. When any sample x is projected to projection axis w, the projection coordinate of x sample can be obtained, that is

$$y = w \bullet x = \sum_{i=1}^{n} a_i y_i k(x_i, x) \tag{10}$$

Therefore, two class samples can be separated through the w projection axis well. For multiclass samples, such as c class samples, we can do as follows: first of all, one class samples are selected from multiclass samples, the other $c - 1$ class samples are used as the second class samples. Then we can get one projection axis. For the above $c - 1$ class samples, the same steps are taken. We can eventually get $c - 1$ projection axes. Hence, when any one sample is projected to $c - 1$ projection axes, we can get $c - 1$ dimension expression features.

In order to meet the real-time requirements, multiclass support vector machine feature extractor is built by a simplified structure. According to Ref. [7], the 4 sub-SVMs are designed for anger, disgust, happiness, sadness, neutral, fear and surprise.

3 Design of Multi-DSP Combined Structure System

The main computation of input object projection is in computing kernel function between new input data and support vector machine, which requires a large amount of multiply-accumulate (MAC) computation. The high speed DAVINCI series DSP chips of TI corporation are used to realize the feature extraction method in this paper. The efficient instruction, proprietary hardware multiplier and efficient parallel structure of DSP can be used to compensate the complexity of algorithm and the cost of operating time. For instance, TMS320DM648 chip has a high performance TMS320C64PLUS + DSP core(900 MHz) system structure [8] and TMS320DM6446 chip has a dual core system structure with the high performance TMS320C64 + DSP(600 MHz) core and the RISC ARM926EJ-S core(300 MHz) [9]. A multi-DSP combined system has been designed according to the serial and parallel characteristics of multiple classifiers structured by four support vector machine. The multi-DSP system structure is shown in Fig. 1.

The system consists mainly of acquisition and preprocessing module, feature extraction module based on support vector machine and classifier and display module. The acquisition and preprocessing module, which is composed of CCD and DM648, is responsible for the image acquisition, face detection to captured images or existing image, face area extraction and face area image being normalized to 30×30 pixel size. The method in Ref. [10] is used for face detection in this paper. The feature extraction module mainly realizes the image feature extraction method based on support vector machine proposed in this paper. This module consists of four TMS320DM648 chips, which are connected with each other. The high-speed point to point LVDS communication is adopted between digital signal processors. The TMS320DM648 #2 chip, which is connected to the TMS320DM648 #1 chip by EMIFA bus, controls the acquisition and preprocessing module and transfers commands and image data between two modules. The classifier and display module is the main processing modules in the system, which consists of a dual core TMS320DM6446 (ARM + DSP) chip with LCD and takes charge of the control and management of the whole system. The ARM core is responsible for the control of entire system and the DSP core is responsible for classification recognition algorithm of the first-order nearest neighbor classifier. The master processor DM6446, which is connected to the host port interface (HPI) of four slave processors in feature extraction module through EMIFA bus, is used to control the slave processors and transmit commands and features data between the master-slave processors.
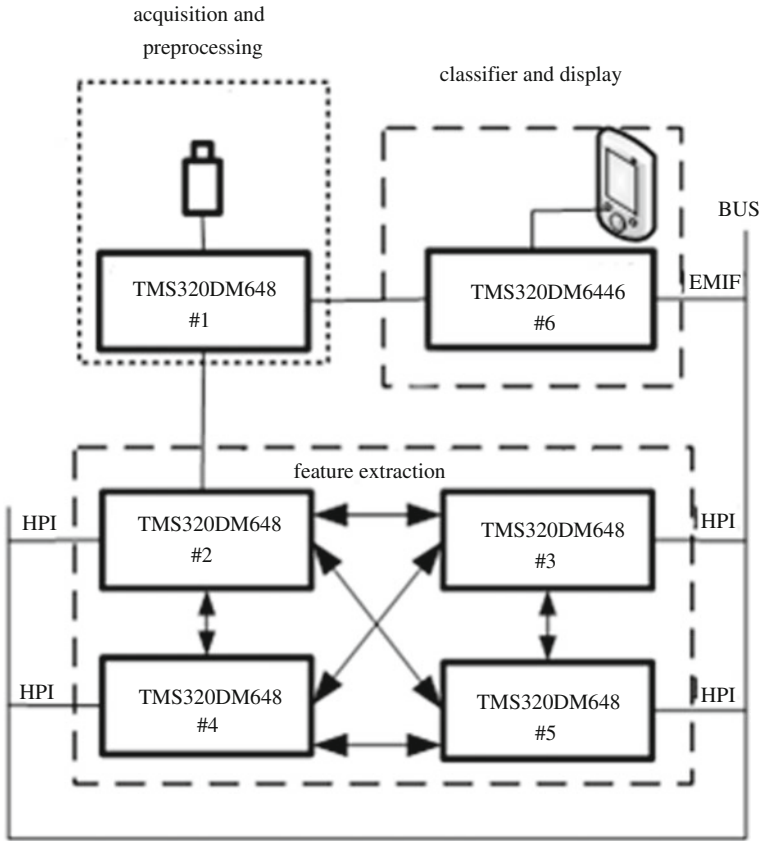
Fig. 1 Structure of multi-DSP system

Each digital signal processor chip has its own peripheral extended DDR2 and FLASH in the system, which supports the high-speed DDR2 and EDMA transmission mode and can meet the requirements of high speed transmission and calculation, and also ensures high-speed parallel ability. The data which is to be processed and analyzed by DSP are stored in DDR2 chips. The algorithm codes and parameters are stored in FLASH chips. The support vector and coefficient table are stored in the extended FLASH of DSP related with support vector machine.

In order to realize real-time image recognition, some special attentions must be paid to the software implementation of algorithm in addition to the above simplified SVM multiclass structure and hardware serial parallel execution, which are as follows. Firstly, simulate algorithm with MATLAB software and improve the efficiency of the algorithm as far as possible, and then generate C program code for TIC6400 by simulation module. Secondly, optimize the algorithm based on support vector machine with linear assembly method, especially for calculation of

kernel function between the new input data and support vector machine, so as to realize parallel pipeline execution of the algorithm. Thirdly, use four threads to implement in each chip, such as HWI, SWI1, SWI2, and IDL ordered from highest to lowest priority. The HWI thread is for hardware events, the SWI1 for implementation of recognition algorithm, the SWI2 for master-slave interactive data update, and the IDL for idle execution respectively.

4 Analysis of Experimental Results

The feature extraction method based on support vector machine is applied to expression recognition in this paper. In order to verify the validity of the presented method, four sub-SVMs are designed and six projection axes are trained out for anger, happiness, disgust, sadness, neutral, fear and surprise. There are 70 dimension expression features projected into this six projection axes for any input sample. In order to meet the real time requirements, four high speed DSP (TMS320DM648) chips with 900 MHz are used to extract facial expression features. At the same time, the high speed LVDS and the EDMA without CPU are used to move data quickly. The recognition process steps in system are as follows:

Step 1: Use DM648 #1 in the acquisition and preprocessing module to detect face for captured images or existing image, and extract the face area image and normalize to 30×30 pixel size.

Step 2: Use 4 DSPs in feature extraction module based on support vector machine to project the preprocessing expression images to six projection axes, and then obtain 70 dimension expression features.

Step 3: Use the DSP core of dual-core DM6446 chip in classifier and display module to recognize expression features with first-order nearest neighboring classifier.

The expression recognition experiments were carried out not only on Yale Face database and Japanese women expression JAFFE database used widely, but also on natural real-time acquiring face expression images. For JAFFE database, we take out 10 samples and 20 samples of each expression respectively, and total training samples are 70 and 140. For Yale Face database, we take 12 samples and 18 samples of each expression respectively, and total training samples are 84 and 126. For real-time captured images, we take 15 samples and 20 samples of each expression respectively, and total training samples are 105 and 140. Some experiments are done repeatedly for the above samples and the results are showed in Table 1. Each line in Table 1 indicates the recognition results in the same kind of feature extraction method with different database, and each column indicates the recognition results on the same database using different feature extraction method.

From Table 1, it can be seen that the feature extraction method based on support vector machine has obvious advantages which mainly come from its own excellent properties, especially good generalization ability to unknown samples.

Table 1 Recognition results for different feature extraction methods

Method	Yale Face (%)	JAFFE (%)	Real-time (%)
PCA	81.24	82.91	75.43
KPCA	81.67	82.91	76.82
LDA	82.63	83.67	81.60
SVM	90.71	91.52	90.61

On the other hand, the projection strategy used in this paper is also one of the reasons. In real-time acquisition case, the recognition performance of PCA, KPCA and LDA method decrease remarkably, which can be due to environment noise and light, and so on. The recognition rate based on support vector machine method has no significant decrease, which may be due to the generalization ability to unknown sample.

5 Conclusion

In order to improve the efficiency and recognition rate of image recognition, the image feature extraction method based on support vector machine is put forward in this paper. This method makes use of many inner excellent properties of support vector machine algorithm, such as the good adaptation ability and generalization ability to unknown samples, which makes the original complex problem simplified, and also improves the problem solving efficiency. This method is also good for improving the system recognition rate and robustness. In this paper, the multi-DSP combined structure is used to implement the proposed method of feature extraction. Besides, the architecture of multiclass support vector machine and the design of hardware and software are optimized in depth to meet the real-time requirements and performance requirements. The results also show the presented method is very effective.

References

1. Yang FU, Hong B (2006) Independent Component Analysis Principle and Application. Tsinghua University Press, Beijing
2. Duda RO, Hart PE, Stork DG (2001) Pattern classification. John Wiley & Sons, New York, 2nd edn
3. Scholkopf B, Smola A, Muller KR (1997) Kernel principal component analysis. In: Gerstner W, Germond A, Hasler M, Nicoud JD (eds), Artifical Neural Networks ICANN 97[C]. Berlin, pp 583–588

4. Mika S, Ratsch G, Weston J, Scholkopf B, Muller KR (1999) Fisher discriminant analysis with kernels. In: Hu YH, Larsen J, Wilson E, Douglas S (eds). Neural Networks for Signal Processing IX[C], IEEE, pp 41–18
5. Roweis S, Saul L (2000) Nolinear dimensionality reduction by locally linear embedding. Science 290(5500):2323–2326
6. Wang Y, Zhou ZH et al (2006) Machine learning and application. TsingHua University Press, Beijing
7. Wang LG, Zhang Y, Gu YF (2005) The research of simplification of structure of multi-class classifier of support vector machine. J Image Graph 10(5):571–574
8. Texas Instruments: TMS320DM648 Digital Media Processor, (2007) http://www.ti.com
9. Sahbi H (2003) Coarse-to-fine support vector machines for hierarchical face detection. PhD Thesis, University of Versailles

Study on the Use of IDAs in Cloud Storage

Honglei Wang and Xuesong Zhang

Abstract Cloud storage is a model of networked online storage based on cloud computing, and provides users with immediate access to a broad range of resources and applications. Although a lot of cloud storage providers adopt encryption to protect costumer data, but users still suspect the security and privacy of their data. The paper analyzed Information dispersal algorithms (IDAs), and proved that it can better address the issues of confidentiality, integrity and availability of data. On this basis, the paper presented a cloud storage system adopting IDAs, and illustrated its key component and the process of writing file in cloud storage.

Keywords Cloud storage · IDAs · Security · Proxy server

1 Introduction

Cloud storage is based on cloud computing, is the extension of Distributed Computing. Cloud storage is somehow the same as cloud computing, they consolidate all storage devices by the features like cluster application, grid technology and distributed file system.

At present, there are several wellknown cloud storage services. Amazon S3 (Simple Storage Service) is an online storage web service offered by Amazon Web Services. It provides storage through web services interfaces (REST, SOAP, and BitTorrent). Google Cloud Storage is a restful online storage web service for storing and accessing data on Google's infrastructure. It is an Infrastructure as a

H. Wang (✉) · X. Zhang
XuZhou College of Industrial Technology, JiangSu, China
e-mail: 98312226@163.com

X. Zhang
e-mail: zxscedar1@163.com

Service (IaaS), comparable to Amazon S3. EMC Atmos is a cloud storage services platform developed by EMC Corporation. Atmos can be deployed as either a hardware appliance or as software in a virtual environment [1].

1.1 Cloud Storage Advantages

Cloud storage is a model of networked online storage where data is stored in virtualized pools of storage which are generally hosted by third parties. Hosting companies operate large data centers, and people who require their data to be hosted buy or lease storage capacity from them. The data center operators, in the background, virtualize the resources according to the requirements of the customer and expose them as storage pools, which the customers can themselves use to store files or data objects. Physically, the resource may span across multiple servers [1].

Cloud storage is made up of many distributed resources, but still acts as one. Cloud storage is highly fault tolerant through redundancy and distribution of data, and highly durable through the creation of versioned copies.

Companies need only pay for the storage they actually use as it is also possible for companies by utilizing actual virtual storage features like thin provisioning. They do not need to install physical storage devices in their own datacenter or offices, but the fact that storage has to be placed anywhere stays the same.

Cloud storage provides users with immediate access to a broad range of resources and applications hosted in the infrastructure of another organization via a web service interface [1].

1.2 Cloud Storage Problem

However, with third party managing the computing resources, off-site data storage raises several security and privacy concerns. Files distributed among a cluster of machines, often span national boundaries. It is not always possible to say under which jurisdiction and data protection laws the out-sourced information falls, and who will consequently be able to access it. Therefore, this usually means data owners lose the control over their data security and privacy. Companies feel nervous about their data, and do not sure whether their data has been adequately protected against theft from attackers both outside and inside the "cloud". Particularly sensitive customer and personal data need to be protected and are subject to heavy constraints that cannot always be matched by current cloud storage solutions. Although a lot of cloud storage providers adopt the most convenient way, encrypting data, in order to protect costumer data, but they usually do the key management themselves. As a consequence, users have no influence on the file encryption process and lose control over who may have access to their data [2]. To ensure confidentiality of data, Customers usually have to select a most reliable

cloud storage supplier, or resort to encrypting their data before sending them to cloud storage device. To ensure availability and integrity of data, Customers normally replicate their data. However these measures have their limitations in a public cloud environment.

Storage system which is based on information dispersal will better address the issues of confidentiality, integrity and availability of data [1]. Current research and commercial implementations storage systems employing IDAs are more suited for static file storage and retrieval. CleverSafe, focuses on storage, using information dispersal to slice TCP/IP packets and store them on a network of local or remote servers in a system called Dispersed Storage. The company's Accesser product slices the data using IDAs [3].

"Information dispersal algorithms (IDAs) have the ability to disperse data in a very secure way across a number of nodes so that if you compromise one node, you won't compromise any data", said Michael Versace, a Wikibon Project partner and analyst. "We're hearing there are a lot of people looking at IDAs as a replacement or an alternative to traditional data encryption" [3].

2 Overview of Information Dispersal Algorithms

Information dispersal algorithms (IDAs)—first proposed by algorithm researcher Michael O. Rabin in 1989—are used to slice data into pieces at the bit level so that when data traverses the network or sits in storage arrays, it is unrecognizable unless accessed by a user/device with the right key. When accessed with the right key, the information is reassembled [3].

2.1 The Main Algorithm

In IDA, a file is split into n slices and a minimum of m slices ($n > m$) are required for reconstructing the original file. A transform matrix of n rows and m columns is used to perform the transformation [4].

Let F be the original file of size N and be a byte array. The bytes in F can be chunked into blocks of m bytes, as in (1):

$$F = (b_1, b_2, b_3, \ldots b_m), (b_{m+1}, b_{m+2}, \ldots b_{2m}) \ldots (b_{N-m+1}, \ldots b_N) \qquad (1)$$

Let A be the transform matrix, as in (2):

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{bmatrix} \qquad (2)$$

Let B be the input file matrix of F and C the output file matrix. Therefore, the following matrix equation can be obtained, as in (3):

$$A \cdot \begin{bmatrix} b_1 & b_{m+1} & \cdots & b_{N-m+1} \\ b_2 & b_{m+2} & \cdots & b_{N-m+2} \\ \vdots & \vdots & \ddots & \vdots \\ b_m & b_{2m} & \cdots & b_N \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1N/m} \\ c_{21} & c_{22} & \cdots & c_{2N/m} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nN/m} \end{bmatrix} \tag{3}$$

To get c_{11}, because c_{11} equals to row 1 of A multiplied by column 1 of B, the following equation can be used, as in (4):

$$c_{11} = a_{11}b_1 + a_{12}b_2 + \cdots + a_{1m}b_m \in GF(2^8) \tag{4}$$

Each row of C corresponds to a slice.

Assume the slices 1 to m to be used for recombination. Let A' be a subset of A of rows 1 to m. Let A^{-1} be the inverse matrix of A', as in (5):

$$A^{-1} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix} \tag{5}$$

For reconstruction, the inverse matrix A^{-1} is applied to the m (out of n) slices to obtain the original data, as in (6):

$$A^{-1} \cdot \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1N/m} \\ c_{21} & c_{22} & \cdots & c_{2N/m} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mN/m} \end{bmatrix} = \begin{bmatrix} b_1 & b_{m+1} & \cdots & b_{N-m+1} \\ b_2 & b_{m+2} & \cdots & b_{N-m+2} \\ \vdots & \vdots & \ddots & \vdots \\ b_m & b_{2m} & \cdots & b_N \end{bmatrix} \tag{6}$$

2.2 Security in Cloud Storage

In Cloud Storage, IDAs can be employed to split files into multiple data slices which will be redundantly stored on several storage nodes. IDAs can also enhance the stored data confidentiality. In a cloud storage system employing IDAs, a adversary who wants to read a file, has to compromise minimum m slice stores or eavesdrops on m slices [4]. He/she also needs to determine which slices logically belong to the file. The adversary will also have to guess the transform matrix of the file and apply the matrix with the correct sequence of the slices. To achieve all these will be very difficult, in practical. The probability of success is very little. Therefore, the storage system employing IDAs, rather than that just employing encryption, can more effectively guarantee the confidentiality of the data.

To further enhance the security of IDAs-based Cloud Storage, the proxy server can optionally encrypt the file slices before sending them to external storage services. The system can be configured to use a randomly generated matrix for every file, instead of using a global matrix for all the files [4].

3 Cloud Storage System Using IDAS

To overcome the security and privacy issues with storing data in cloud computing network as they were discussed in previous sections, we propose a system architecture adopting IDAs for securing off-site data storage, depicted in Fig. 1.

3.1 Proxy Server

The key component of this system is a proxy server which is responsible for integrating the external storage services from the Internet, offering the new resources to the client computers on the intranet, and securing all data transfers as soon as they leave the trusted enterprise network zone [2].

The proposed proxy server can be a Linux based system, and usually be located within the trusted zone of a company's intranet. One of the major goals is to seamlessly integrate the external cloud storage services into an employee's desktop work-space using the proxy server as a mediator [2]. Therefore, the user no longer has to add any additional software components for storing files in Cloud

Fig. 1 System architecture
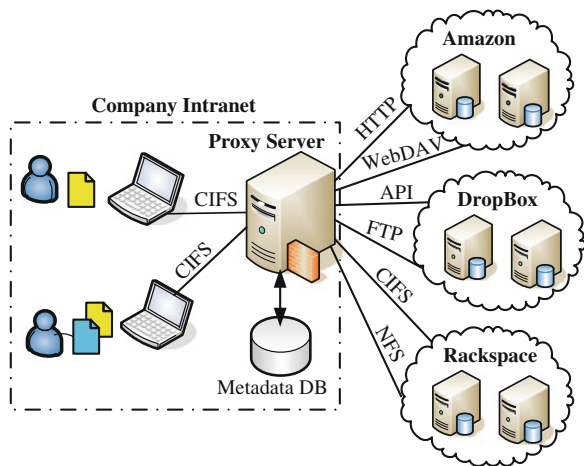
Storage. Through Common Internet File System (CIFS), the user can employ the additional storage resources offered by the proxy server on the client computer. CIFS is a protocol that lets programs make requests for files and services on remote computers on the Internet [3]. It uses the client/server programming model, and seems to be the best choice in the mostly Windows dominated world of office computers [2].

For the purpose of starting the dispersion and encryption algorithms on the proxy server, a customized file system is necessary that enables users to "overwrite" the standard file system operations. The famous Filesystem in Userspace (FUSE) will be adopted here to achieve the necessary functionality [2]. FUSE is a loadable kernel module for Unix-like computer operating systems that lets non-privileged users create their own file systems without editing kernel code [1].

3.2 Write Operations

The process of writing file in Cloud Storage will be carried out as follows, depicted in Fig. 2: the user usually a company employee copies a file to a desired folder on the network drive. This file will be cached on the proxy server, almost at the same time the proxy server will generate a special matrix for the file randomly. Then the proxy server will use the matrix to transform the file into multiple slices. The resulting data slices will now be redundantly stored either on locally attached storage, e.g. on a NAS, or on one of the online cloud drives using the protocol adapter [2]. At last, the data fragments will be encrypted additionally to enhance the confidentiality of information before leaving the intranet [5].
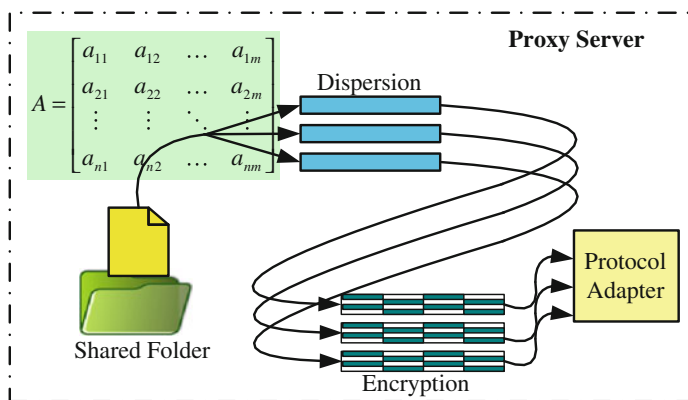

Fig. 2 Writing process

3.3 Metadata DB

In order to maintain the mapping of file to file slices, class FileInfo and class FileSlice are used as shown in Fig. 3. Class FileInfo contains properties such as IDA matrix used to slice the file, namespaces of slice stores, as well as general file attributes and so on [2]. Class FileSlice contains information of one file slice. A file can be divided into a number of file slices. All file instances are stored in metadata database. During the whole writing process additional information and metadata belonging to the out-sourced file will be stored into the database which allows the cached file to be deleted from the proxy server after the storage procedure was completed successfully [6].

By employing metadata database, proxy server can manage slice store registration and removal, as well as keeping track of the health and availability of all the slice stores [4]. The database is going to be a MySQL db, which should be replicated, distributed, and protected against attacks and failures according to best practices in the field of database security [2].

4 Performance

We have carried out a preliminary performance test for IDAS Storage System read and write operations of different file sizes. We also compared the IDAS Storage System performance with that of NFS for benchmarking. Our test is carried out on a VMware ESXi 5.0 private cloud, and is based on 4 active slice stores. The NFS client and server used are based on Ubuntu 10.10. The registry server, slice stores and NFS servers are running on VMs configured with 1 GB of memory and 1 virtual CPU with no resource reservation.

From the analysis of performance data, it can be seen that read and write times increase linearly with the size of the file. Figure 4 shows that the IDAS Storage System write performance is on a par with that of NFS write. In order to improve the IDAS Storage System read performance, a read-ahead caching has been employed in our test. Therefore, the IDAS Storage System read is not significantly slower than NFS read as shown in Fig. 4.
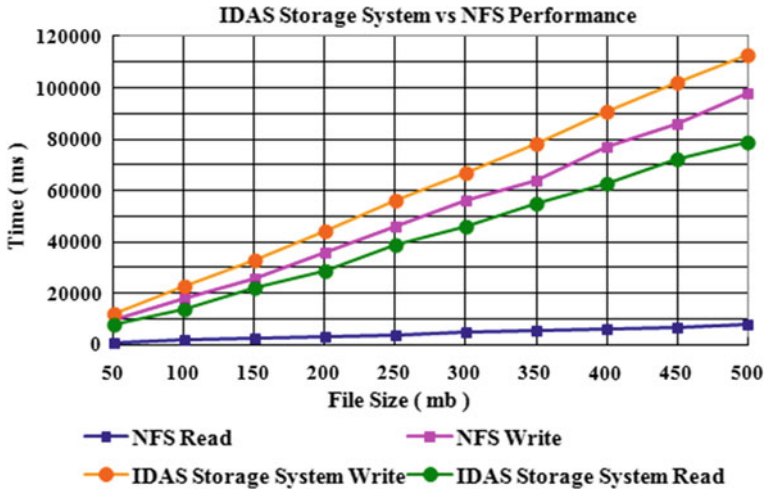
Fig. 3 Metadata DB
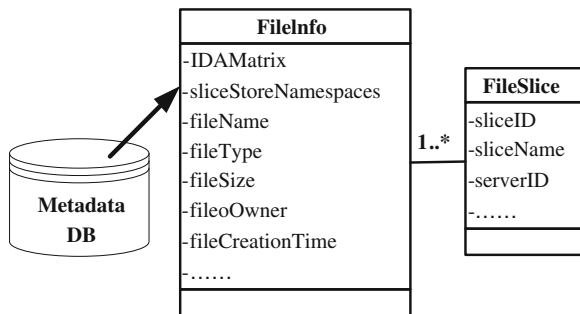
Fig. 4 IDAS storage system VS NFS performance

5 Conclusion

In this paper, we presented a system architecture that provides data security, integrity and availability suited for deployment in a public cloud environment. The system's core component is a proxy server which is responsible for data distribution via IDAs. In order to further enhance the data security, the proxy server will generate a special matrix for each file randomly, and will encrypt file slices before sending them to cloud storage services. Due to this fact that all operations are performed in the trusted local intranet, the usual security concerns and privacy issues with common cloud storage services should be mitigated.

References

1. Hou Q, Wu Y, Zheng W, Yang G (2011) A method on protection of user data privacy in cloud storage platform. J Comput Res Develop 48(7):1146–1154
2. Seiger R, GroB S, Schill A (2011) SecCSIE: a secure cloud storage integrator for enterprises, IEEE conference on commerce and enterprise computing, pp 252–255
3. Jiyi W, Jianqing F, Lingdi P, Qi X (2011) Study on the P2P cloud storage system. ACTA Electronica Sinica 139(5):1100–1106
4. Mar KK (2011) Secured virtual diffused file system for the cloud. 6th international conference on internet technology and secured transactions, United Arab Emirates, pp 116–121
5. Shixin L, Fengmei L, Ren C (2011) A hierarchy attribute-based access control model for cloud storage. International conference on machine learning and cybernetics, China, pp 1146–1150
6. Xu L, Hu J, Mkandawire S, Jiang H SHHC: a scalable hybrid hash cluster for cloud backup services in data centers. 31st International conference on distributed computing systems workshops, pp 61–65

Influence of Potential Parameters on the Melting Temperature of MgSiO$_3$ Perovskite

Qiong Chen

Abstract The melting temperatures MgSiO$_3$ perovskite have been calculated in previous studies by using MD simulation, but considerable discrepancy of melting temperature exists between these simulations. In this contribution, comparisons of potential energy curves are performed to explain the discrepancy. To further investigate the influence of the interaction potential parameters on the MD simulation result, a new set of potential parameters is developed based on combining two fitting potential parameters of previous studies, and is applied in the present study. The melting temperatures are calculated, and also compare with those derived from previous studies.

Keywords Potential parameters · Molecular dynamics · Melting temperature

1 Introduction

The most common mineral in the Earth's lower mantle is MgSiO$_3$ with the perovskite structure. An understanding of its thermodynamic properties is important to describe the Earth's evolution. Therefore the physical properties of MgSiO$_3$ perovskite under extreme pressure and temperature conditions have attracted the attention of geophysicists [1–3].

The subject of MgSiO$_3$ perovskite melting has not yet been well understood. The experimental data on material properties at high pressures is still limited, especially under the conditions of both high temperature and high pressure. Computer modeling is particularly useful and powerful for problems that may be inaccessible to direct experiment study, such as extremes in temperature and

Q. Chen (✉)
College of Electrical Engineering, Northwest University for Nationalities,
LanZhou 730124, China
e-mail: chenqiong103@163.com

pressure prevailing deep in the Earth. The prediction of melting by molecular dynamics simulations has been found to be rather extensively applicable, and, furthermore it allows tracking of the physical properties of the atoms not only by global averages but also locally. However, a long-standing disagreement exists concerning the melting of $MgSiO_3$ perovskite. The melting temperatures simulated on a perfect lattice system [4] appear to be rather high compared to the melting temperature observed in experiments (ZB [5], HJ [6], KJ [7] and SH [8]). The experimental melting temperatures are significantly lower than the values calculated by Chaplot using a system with 1 % vacancies [4], and lower than the values reported in the Refs. [9] and [10]. Alfè et al. [11] have enumerated possible reasons for the discrepancy. For example, the sample may have been subjected to non-hydrostatic or thermal stress (leading to crystal failure for $T < T_m$ [12]). The method of locating the melting point in diamond anvil cell (DAC) experiment at $\sim 4,000$ K may be questioned, because the eye-evident surface texture change is usually used as the melting criterion to the melting of sample heated by laser. Aguado [13] considers that the premelting effect in the melting process is a possible explanation for the difference between experiment and simulation. The above discussions indicate that melting may be a complicated phenomenon.

The previous studies [9, 10] have chosen the same pair-potential model, but with different potential parameters. The melting curve of Ref. [10] is in better agreement with the experimental data. It is important to know the influence of the interaction potential parameters on simulation results, but there is no such study to the best of our knowledge. In this contribution, the authors investigate the influence of the potential parameters on the simulation results of $MgSiO_3$ perovskite. First of all, comparisons of potential energy curves are made to explain the difference of melting temperature that exists in previous studies [9, 10]. The rigid-ion two-body potential [11], which is another form of potential used in simulating melting of $MgSiO_3$ perovskite, is also compared with the pair potential energy curve. To further study the influence of the interatomic potential parameters on the MD simulation result, a new set of potential parameters is developed, based on combining two fitting potential parameters of Refs. [9, 10]. Finally, the new set of parameters is employed to simulate the melting curve of $MgSiO_3$ perovskite.

2 Method

2.1 Simulation Technique

MD simulation is a well-established technique to determine the properties of materials. This method has been applied extensively in previous works [15–18]. Normally, as is often the case in MD calculation, periodic boundary conditions (PBC) are applied. The electrostatic interactions are evaluated both in real and reciprocal space according to the Ewald's method [19].

The present calculations are performed with the shell-dynamo code [20]. Simulations in NTP (constant N is the number of particles, T is the temperature, and P is the pressure) ensemble [21] are applied. Simulation runs are carried out with 360 particules (72 Mg, 72 Si, 216 O atoms). The results of molecular dynamics simulations in the NTP ensemble with chosen model of the interatomic interaction depend on, apart from the initial arrangement of particles, the size of the time step (Δt), the number of particles (N), and the temperature of windows (Twin). The influence of these parameters was studied by carrying out test runs at various temperatures and pressures. The correct results could be obtained with $\Delta t = 0.001$ ps, Twin $= 23$ K. Usually, the system was re-equilibrated at the required temperature and pressure for 3 ps and then temperature, pressure and volume were measured over a further 6 ps. Typically, the actual temperature differed from the required temperature by less than 1 K, the actual pressure differed from the required pressure by less than 1 MPa, and the discrepancy in the volume was less than 0.002 cm^3/mol.

The constant-pressure heat capacity Cp and thermal expansivity α_P are calculated from the following definitions.

$$C_P = \left(\frac{\partial H}{\partial T}\right)_P \tag{1}$$

$$\alpha_P = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_P \tag{2}$$

Here, H, E and V are enthalpy, pressure and volume, respectively.

2.2 Interaction Potential

In this study, the interaction potential is pair potential, and has the following form:

$$U(r_{ij}) = \frac{Z_i Z_j e^2}{r_{ij}} + A_{ij}\exp\left(-\frac{r_{ij}}{\rho_{ij}}\right) - \frac{C_{ij}}{r_{ij}^6} \tag{3}$$

where the first part is the long-range Coulomb term and the rest are the short-range terms in the form of a Buckingham [22] potential. The Buckingham potential is a rather traditional model that has been performed sufficiently well [23–26] and, therefore is widely used for modeling of various oxides. The advantages and shortcomings of this kind of model are known from Ref. [27]. Here Zi or Zj is an effective charge, e is the electronic unit charge, rij is the interatomic distance between atoms i and j, Aij and ρij are the parameters for the repulsive interaction, and Cij is the van der Waals constant.

3 Results and Discussion

3.1 Comparison of Potential Energy Curve

Figure 1 shows a comparison of potential energy curves. The pair potential energy curve of Ref. [9] is below that of Ref. [10], and the interatomic interaction of Ref. [9] is stronger than that of Ref. [10], therefore the simulated melting temperatures of Ref. [9] are higher than that of Ref. [10]. This indicates that the potential parameters have an important influence on the MD simulations; In the study of Ref. [14], the rigid two-body potential is adopted to study the melting of $MgSiO_3$ perovskite. However, the melting temperature of study of Ref. [10] was closer to the experiment data [5] than that of Ref. [14] at 30 GPa. This discrepancy may be explained by comparing energy curve of the pair potential with that of rigid two-body potential. As shown in Fig. 1, the study of Ref. [14] has a lower potential energy curve than the study of Ref. [10], and the interatomic interaction of Ref. [14] is stronger than that of Ref. [10]. Because of this, $MgSiO_3$ perovskite is relatively easier to melt by using the potential model in Ref. [10] than that of Ref. [14]. For the melting of $MgSiO_3$ perovskite, the rigid ion two-body potential model is not preferable to the pair potential model which is used in the present work.

3.2 Derivation of the New of Potential Parameters

In the crystal structure of $MgSiO_3$ perovskite, there are two distinct oxygen sites at octahedral corners. Silicones lie in the octahedral centers. Each Si atom is octahedrally coordinated by O atoms. Magnesiums are isolated, compare to silicones and oxygens. The octahedral structure is rather stable. Consequently, the

Fig. 1 The potential energy as a function of ion–ion separation for the $MgSiO_3$ perovskite. The dash dot line represents potential energy curve of of Matsui et al. [14], the *dots* represent the potential energy curve of the study of Ref. [10] and the *dashes* represent the potential energy curve of the study of Ref. [9]
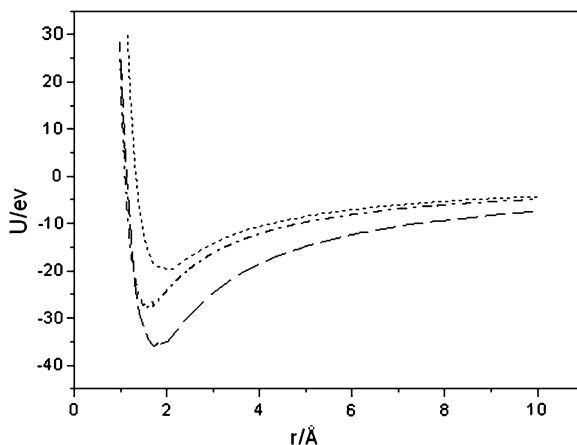
Si–O and O–O interaction may be more important for the properties of MgSiO$_3$ perovskite than the Mg–O interaction. To further investigate the influence of potential parameters on the MD simulation, the authors design a new set of potential parameters. In the new set of potential parameters, the Si–O and O–O potential parameters, shell charges and spring constants are kept at the values used in the study of Ref. [10], the Mg–O potential parameters are taken from the study of Ref. [9]. The method that transforming the potential parameters of one system into another similar system is available in the other literatures for example, in the study of Ref. [28], the potential parameters representing short-range interactions between Al$_{3+}$ and O$_{2-}$ in Al$_2$O$_3$ were used in the MD stimulation of MgAl$_2$O$_4$. As shown in Ref. [29], the technique generated a set of potential parameters for ZnAl$_2$O$_4$ and ZnGa$_2$O$_4$: The Oxygen–Oxygen short range potential were taken from the work of Bush et al. [30]; Zn–O potential were fitted to the hexagonal polymorph of ZnO (wurtzite); and the Al–O and Ga–O potential parameters were determined by fitting to α-Al$_2$O$_3$/ZnAl$_2$O$_4$ and Ga$_2$O$_3$/ZnGa$_2$O$_4$ respectively. Consequently, the potential parameters used in the present study are derived using this technique, the parameters are: AMg–O = 1,041.435 eV; ρMg–O = 0.2673 Å; ASi–O = 7,363.45 eV; ρSi–O = 0.1900 Å; AO–O = 1,621.68 eV; ρO–O = 0.3000 Å; CO–O = 1621.68 eV Å6; ZMg = +1.565; ZSi = +2.329; Zo = −1.298.
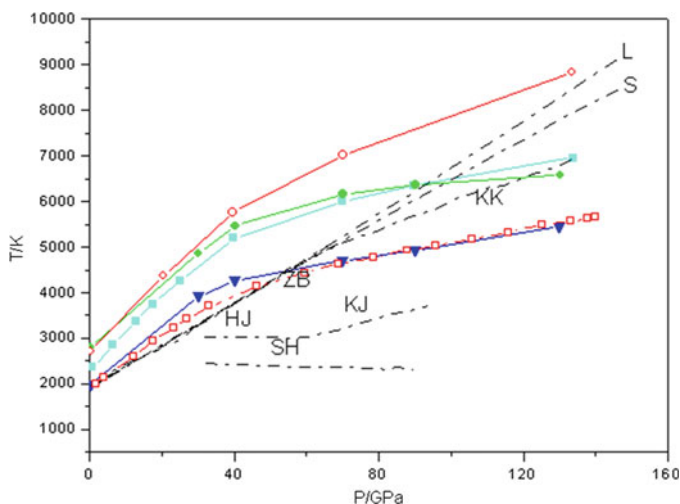
3.3 Melting Curve of MgSiO$_3$ Perovskite

Melting is probably the most familiar first-order transition; it covers the widest range of pressures and temperature. It exhibits discontinuities in the first derivatives of the volume, free energy and entropy. The changes of these physical properties are often used for identification of melting. In the MD simulations, melting can be identified by a sudden change in the atomic volume. The corresponding temperature is the melting temperature under given pressure. In the present simulation, at 0.1 MPa, the sudden change of the calculated volume occurs at a temperature of about 2,407 K. Note that simulated melting temperature is higher than the experiment melting temperature. The reason is that MD simulations in free space are usually known to exhibit high hysteresis especially at low pressure [31]. This prevents us from judging how well the experimental value of melting temperature at 0.1 MPa pressure is reproduced. Cohen and Gong [31] noted the high hysteresis at 0.1 MPa; this hysteresis decreases with the increasing pressure. The high simulated temperature is also the result of the existence of a substantial metastable overheating. Overheating of a crystalline solid occurs when the long-range order of the crystalline structure is maintained up to certain temperature above the equilibrium melting temperature. According to the modern theory of melting, melting is a dynamic process and melting temperatures of a crystal can be modified by certain melting mechanisms [32], the overheating melting of some crystalline solids at ambient pressure can be written as $\theta = T/Tm - 1$, where Tm is the conventional

melting temperature for the bulk. This melting process results from temperature increase rate, is not very dependent on pressure, and can be applied to modify the melting temperatures for overheating of crystal. Luo et al. [33] have investigated the systematic of nucleation energy barrier (β) for elements and compounds, and the corresponding overheating (θ) as a function of heating rate (Q). They pointed out that significant overheating is achievable via ultrafast heating, and demonstrated that the degree of overheating achieved in shock-wave loading and intense laser irradiation, as well as in molecular dynamics simulations (Q $\sim$ 1,012 K/S), agrees with the established
θ-β-Q systematic based on undercooling experiments and homogeneous nucleation theory. Crystalline solids can be superheated by (0.05–0.35) Tm and (0.08–0.43) Tm at 1 and 1,012 K/S, respectively [34]. According to catastrophic melting [35, 36] and homogeneous nucleation [32, 37] theories, the overheating temperature is about 21.2 % higher than normal equilibrium melting point at room pressure. Therefore, in this work the overheating temperature of MgSiO$_3$ perovskite is assumed to be 0.212 Tm in the whole pressure–temperature range. This value yields the melting curve of MgSiO$_3$ perovskite (shown in Fig. 2).

In Fig. 2, the present melting curve of MgSiO$_3$ perovskite is closer to with experimental data reported by ZB [5] than the study of Ref. [9]. Meanwhile, the

Fig. 2 The melting curve of MgSiO$_3$ perovskite. The melting temperatures are significantly lower for the present study (*closed triangles*) than that for the perfect lattice system [4] (*open circles*), the system with 1 % vacancies [4] (*closed squares*) and the simulation [9] (*closed circles*). The melting temperatures are close to simulation [10] (*open squares*). Comparison is made with the experimental data [5] up to 62.5 GPa and its extrapolations (L,S,KK) to higher pressures using different melting relations as reported in ZB and also with earlier laser heating experiment(HJ [6],KJ [7], SH [8], (SH indicates a lower bound only)). The present MD simulation is closer to the experimental data by ZB. The data below 25 GPa are from Refs. [32, 33]

present melting curve has a slight difference from that in the study of Ref. [10]. Part of the reason may be that the Si–O and O–O potential parameters used in present work are same with those used in the study of Ref. [10]. This indicates that the Mg–O interaction parameters have smaller influence on the melting of $MgSiO_3$ perovskite than the Si–O and O–O interaction parameters.

4 Summary

In this paper, comparison of potential energy curves was performed. The data not only provided a possible explanation for the discrepancy of melting temperature between previous studies, but also gave a support for using the pair potential of the present simulation. A new set of potential parameters based on adopting suitable empirical potential parameters was used to simulate the thermodynamic properties of $MgSiO_3$ perovskite. The resulting molar volumes, equation of state, constant-pressure heating capacity and constant-pressure thermal expansivity are close to the experimental data. It was found that the Si–O and O–O interaction is more important than Mg–O interaction for the melting of $MgSiO_3$ perovskite.

References

 1. Oganov AR, Brodholt JP, Price GD (2001) Earth planet. Sci Lett 184:555
 2. Fiquet G, Dewaele A, Andrault D, Kunz M (2000) Geophys Res Lett 27:21
 3. Shim SH, Duffy TS, Shen GY (2001) Science 293:2437
 4. Chaplot SL, Choudhury N, Rao KR (1998) Am Miner 83:937
 5. Zerr A, Boehler R Science (1993) 262:553
 6. Heinz DL, Jeanloz R (1987) Geophys J Res 92:11437
 7. Knittle E, Jeanloz R (1989) Geophys Res Lett 16:421
 8. Sweeney JS, Heinz DL (1993) Geophys Res Lett 20:855
 9. Liu ZJ, Cheng XL, Yang XD, Zhang H, Cai LC (2006) Chin Phys 15:1009
10. Liu ZJ, Sun XW, Yang XD, Cheng XL, Guo YD (2006) Chin J Chem Phys 19:311
11. Alfe D (2004) J Phys Condens Matter 16:S973
12. Belonoshko AB, Dubrovinski LS (1997) Am Mineral 82:441
13. Aguado A, Madden PA (2005) Phys Rev Lett 94:068501
14. Matsui M, Price GD (1991) Nature 351:735
15. Gu YK, Qi YH, Qin YJ (2003) Chin J Chem Phys 16:385
16. Wei YH, Ma XL, Zhang GY Hou Q, Wang H.C, Sun YS (2004) Chin J Chem Phys 17:443
17. Liu ZJ, Cheng XL, Chen XR, Zhang FP (2005) Chin J Chem Phys 18:193
18. Liu ZJ, Cheng XL, Zhang FP (2006) Chin J Chem Phys 19:65
19. Ewald PP (1921) Annal Physik 64:253
20. Fincham D (1994) J Mol Graphic 12:29
21. Fincham D (1994) J Mol Graphic 12:29
22. Catlow CRA (1989) J Chem Soc Faraday Trans 2(85):335

23. Belonoshko AB (1994) Geochim Cosmochim Atca 58:4039
24. Belonoshko AB, Dubrovinski LS (1996) Am Mineral 81:303
25. Matsui M (1996) Phys Chem Miner 23:345
26. Kuklju MM (2000) J Phys: Condens Matter 12:2953
27. Kramer GJ, Farragher NP, van Beest BWH (1991) Phys Rev B 43:5068
28. Catlow CRA (1985) J Phys C: Solid State Phys 18:1149
29. Pandey R, Gale JD, Sampth SK, Recio (1995) JM J Am Ceram Soc 82:125
30. Bush TS, Gale JD, Catlow CRA, Battle PD (1994) J Mater Chem 4:83
31. Cohen RE, Gong Z (1994) Phys Rev B 50:12301
32. Lu K, Li Y (1998) Phys Rev Lett 80:4474
33. Luo SN (1985) Ahrens TJ Phys Rev A 31:1695
34. Luo SN, Ahrens TJ (2005) Appl Phys Rev Lett 94:068501
35. Fecht HJ, Johnson WL (1983) Nature 334:50
36. Tallon JL (1989) Nature 342:658
37. Rethfeld B, Sokolowski-Tinten K, von der linde D, Anisimow SI (2002) Phys Rev B 65:092103

Research on Attack Graph Generation for Network Security Situation

Yanbo Wang, Huiqiang Wang, Chao Zhao, Yushu Zhang and Ming Yu

Abstract Attack graph generation method based on network security situation is presented. Attack graph technique bases attack graph on the target network and the attack model. Generally, attack path is shown that the attacker uses vulnerability of target network to carry out network attack by graph structure. Attribute attack graph generation method based on breadth-first is put forward, which during the process of the attack graph generation solve the problem of circle path and combination explosion, the different scale of simulation experiment shows that the research results can found in time and make up for security problems existing in the network system, effectively improve the survivability of the network system, so as to improve the ability that network system deals with all kinds of sudden attack.

Keywords Attack graph · Attack path · Vulnerability · Circle

1 Introduction

With the development of Internet technology, the network has been used in various fields, such as government, army, energy and spaceflight, etc. The network brings great convenience to people's life and work; however, there are a lot of network attack behavior, such as viruses, Trojan, denial of service and Bot-Net, etc., which seriously influence economic security and defense security. At the same time, there

Y. Wang (✉) · H. Wang · C. Zhao · Y. Zhang
College of Computer Science and Technology Harbin Engineering University,
Harbin, China
e-mail: Heu_wangyanbo@126.com

M. Yu
Information and Computer Engineering College Northeast Forestry University,
Harbin, China

is always some network security vulnerabilities, in which certain relationship may exist, that is when a bug is utilized successfully, it may leads to another bug used. In order to completely find out the relationship among all attack incidents, the most effective way is that through simulating the attack process that assailant attacks network which exists security vulnerabilities, find out the attack path that attacker accomplishes the target, and at the same time, these paths are shown as graph, that is network attack graph, which has stronger description ability of network attack process and the more wide application range compared to attack tree or Petri nets [1–3].

According to the network state and vulnerability information, attack graph technology analyses the vulnerability sequence used by attacker, with that a directed graph is constructed, namely attack graph. Attack graph technology firstly build model for network, which indicates actual state of target network, such as host configuration information, host vulnerability information, network topology information, network configuration information, etc. And then it analysis through the attack mode the steps may be taken to achieve target with multiple vulnerability by attacker. Attack graph technology with the form of graph shows the network vulnerability relationship, which makes network security manager observe intuitively vulnerability existing in the network and all attack paths in the network, facilitate network security protection for network security manager [4–6].

2 Attack Graph and Related Concept

Definition 1 Attack Graph $AG = (A_0 \cup A_d, T, E)$ is a directed graph, where A_0 represents a set of initial nodes, and A_d represents a set of reachable nodes. T denotes a set of atomic attack nodes; E is a set of directed edges. AG satisfies the following constraint [7–9]:

(1) $E \subset (T \times A_d) \cup ((A_0 \cup A_d) \times T)$: the relationship between two nodes in attack graph only includes $A_0 \rightarrow T$, $A_d \rightarrow T$ and $T \rightarrow A_d$, where $T \rightarrow A_d$ denotes consequence edges of atomic attack; $A_0 \rightarrow T$, $A_d \rightarrow T$ denotes precondition edges of atomic attack.
(2) $\forall \tau \in T$, $\mathrm{Pr}\,e(\tau)$ represents a set of τ's father nodes, $\mathrm{P}\,ost(\tau)$ represents a set of τ's child nodes. The "and" relationship exists among father nodes, and $\wedge \mathrm{Pr}\,e(\tau) \Rightarrow \wedge \mathrm{P}\,ost(\tau)$ denotes when all the precondition of atomic attack is satisfied, the atomic attack could happen.
(3) $\forall a \in A_d$, $\mathrm{P}\,arent(a)$ represents a set of a's father nodes, The "or" relationship exists among father nodes, and $\vee \mathrm{P}\,arent(a) \Rightarrow a$ denotes any atomic attack in $\mathrm{P}\,arent(a)$ is implemented successfully, which makes a is satisfied.

Shown by the above definition, attack graph contains attribute node and atomic attack node. Attribute node represents the condition of target network attributes and the attacker ability; atomic attack node represents an attack that the attacker uses

some vulnerability. Figure 1 shows a attack graph of the target network. The target network contains two hosts, Number one and two; the attacker connected to the network with a malicious host outside of the target network, the malicious host Number is zero. In Fig. 1, the text characters mean the attributes of target network and attacker, detailed description shown in Table 1, ellipse denotes atomic attack, detailed description seen in Table 2.

Definition 2 In this paper, vulnerability is shown as $(hostid, vid, range, type, service, result)$. $hostid$ is hostname where the vulnerability exists, vid is id in CVE vulnerability Database, $range$ is use range of vulnerability, $type$ is the type of vulnerability, $service$ is the corresponding service name, $result$ is the obtained result after vulnerability is utilized successfully.

Definition 3 In attack graph, state node is shown as $(sid, hostid, vid, result, p)$. where, sid is state node number, $hostid$ is hostname whose safety elements have changed in the network state, vid is vulnerability number used to reach the network state, $result$ is obtained result by vulnerability, p is probability that attacker successfully reach the network state.

Definition 4 Aggressive behavior is shown as $(src\_host, dst\_host, vid)$. src_host denotes attack host id, dst_host is attacked host id, vid is used vulnerability number.

Definition 5 Vulnerability using rule is formal description for a kind of precondition of aggressive behavior and the obtained results from successfully used vulnerability. $c_1 \wedge c_2 \wedge \ldots \wedge c_n$ is precondition, and c_0 is result, so the rule can be expressed as: $c_1 \wedge c_2 \wedge \ldots \wedge c_n \Rightarrow c_0$.

Fig. 1 Example of attack graph
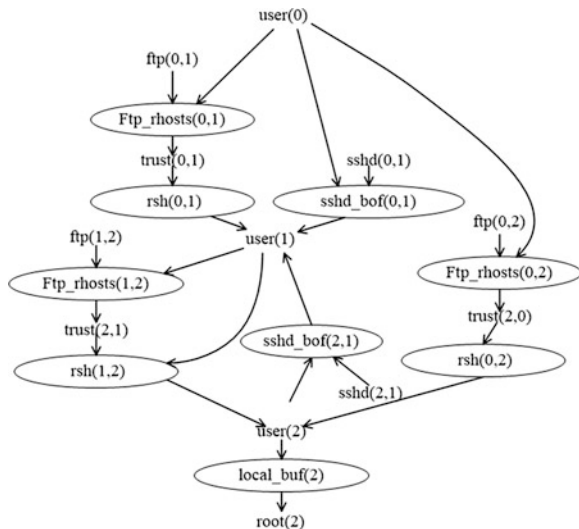
Table 1 The attributes in the example of attack graph

Attribute	Description
ftp(a,b)	The service ftpd on host b is accessible from host a
Trust(a,b)	Host a trusts host b
User(a)	Attacker has the user privilege on host a
Root(a)	Attacker has the root privilege on host a

Table 2 Atomic attacks in the example of attack graph

Atomic attack	Description
Ftp_rhosts(a,b)	Attack establishes a remote login trust relationship from host a to host b via the ftp_rhosts
sshd_bof(a,b)	Attacker gains the user privilege on host b from host a using a remote buffer overflow attack on the sshd vulnerability of host b
local_bof(a)	Attacker achieves the root privilege on host a using a local buffer overflow attack on host a

Definition 6 Attack node is shown as: $(\Pr e(e), e, P ost(e))$, e represents certain an aggressive behavior, $\Pr e(e)$ is a set of precondition that e is successfully carried out, $P ost(e)$ is a set of result that e is successfully carried out. It is an instance that aggressive behavior and vulnerability utilizes rules.

Definition 7 $path = \bot \rightarrow \tau_1 \rightarrow \tau_2 \rightarrow \ldots \rightarrow \tau_l$ is atomic attack sequence in attack graph AG, where, $\bot$ represents the starting identifier of sequence, $\tau_i \in T(1 \le i \le l)$ denotes atomic attack. If the sequence satisfies the following constraints, we will say it is an attack path of A_c [10]

(1) the precondition of any atomic attack latter is consequences or initial attribute of atomic attack in front. $\forall a \in \Pr e(\tau_i), \quad a \in \bigcup_{k=1}^{i-1} P ost(\tau_k) \cup A_0$, where, $\tau_i \in T, \quad 1 \le i \le l$;
(2) the result of any atomic attack in front is precondition of atomic attack latter, $\exists b \in P ost(\tau_i), \quad a \in \bigcup_{k=i+1}^{l} \Pr e(\tau_k)$, where, $\tau_i \in T, \quad 1 \le i \le l$;
(3) there are non-null intersection between the result set generated from the last atomic attack and determinant attribute set A_c $P ost(\tau_l) \cap A_c \ne \emptyset$.

In the attack graph example shown in Fig. 1, $A_c = \{user(1)\}$, the attack sequence $path_1 = \bot \rightarrow Ftp\_rhost(0, 1) \rightarrow rsh(0, 1)$ and $path_2 = \bot \rightarrow sshd\_bof(0, 1)$ is two attack paths of A_c. $path_3 = \bot \rightarrow Ftp\_rhost(0, 1) \rightarrow rsh(0, 1) \rightarrow Ftp\_rhost(1, 2) \rightarrow rsh(1, 2) \rightarrow sshd\_bof(2, 1)$ is also an attack path of A_c, which has a circle. But during the actual attack process, the attack path will not happen, because the attacker will not obtain the pre-existing ability. So in the above example, the attacker obtain permission of user in host1, it will not repeatedly obtain the permission of use through host2. But we can't directly delete

Fig. 2 Example of
equivalent attack paths
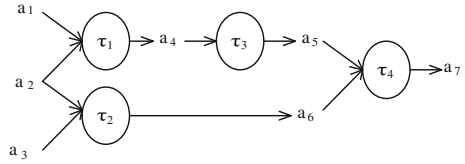
$sshd\_bof(2,1)$, otherwise the reasonable path $path_4 = \bot \rightarrow Ftp\_rhost(0,2) \rightarrow rsh(0,2) \rightarrow sshd\_bof(2,1)$ will be deleted. Attack graph contains circle, which significantly increased the complexity of the attack graph analysis.

Definition 8 $path = \bot \rightarrow \tau_1 \rightarrow \tau_2 \rightarrow \ldots \rightarrow \tau_l$ is an attack path in attack graph AG, if the attack path don't have circle, that is $\forall \tau_i, \tau_j \in T (1 \le i \le j \le l)$ and $\Pr e(\tau_i) \cap \bigcup_{j=i+1}^{l} Post(\tau_j) = \emptyset$, it says the attack path is effective attack path.

Among effective attack paths, the precondition of any atomic attack is not the consequences of atomic attack latter, which reflects the real path that attackers may adopt. To accurately identify the intention of attacker, we must identify all the effective attack path.

Research found that generally in attack graph there are a lot of effective attack path contains exactly the same atomic attack, although these attack sequence is different, but reflect the dependency among the same atomic attack. In order to simplify and easy to understand, we define this kind of effective attack path is equivalent. Shown in Fig. 2. τ_3 depends on τ_1, but there is no dependence between τ_2 and τ_1, τ_3, so $\bot \rightarrow \tau_1 \rightarrow \tau_2 \rightarrow \tau_3 \rightarrow \tau_4$, $\bot \rightarrow \tau_1 \rightarrow \tau_3 \rightarrow \tau_2 \rightarrow \tau_4$ and $\bot \rightarrow \tau_2 \rightarrow \tau_1 \rightarrow \tau_3 \rightarrow \tau_4$ are equivalent effective attack path.

Definition 9 $path = \bot \rightarrow \tau_1 \rightarrow \tau_2 \rightarrow \ldots \rightarrow \tau_l$ is an effective attack path in attack graph AG. $n \in N$ is constant, if the length of attack path $l \le n$, it says this path is n-effective attack path [11].

For AG, if $a \in Post(\tau_l)$, we will say $path^{(i)}(a) = \bot \rightarrow \tau_1 \rightarrow \tau_2 \rightarrow \ldots \rightarrow \tau_l$ is ith attack path to a. $PATHS(a) = \{path^{(i)}(a)\}$ denotes set of attack paths to a. If $\tau = \tau_m$, it says that $path^{(j)}(a) = \bot \rightarrow \tau_1 \rightarrow \tau_2 \rightarrow \ldots \rightarrow \tau_m$ is jth attack path to τ, $PATHS(\tau) = \{path^{(j)}(\tau)\}$ represents set of attack paths to τ.

3 Attack Graph Generation Algorithm

Algorithm：Attack graph generated algorithm based n-effective attack path

Input：A_0 , T_0 denotes available vulnerability set , N is search range

Output：AG

01 $A_d \leftarrow \varnothing$, $T \leftarrow \varnothing$, $E \leftarrow \varnothing$

02 $depth \leftarrow 0$, $A_0.depth \leftarrow depth$, $waiting\_queue \leftarrow A_0$

03 $while \quad waiting\_queue \neq \varnothing \quad and \quad depth \leq 2N$

04   $do \quad a \leftarrow waiting_{queue}.pop(\ )$

05    $if \ a.depth > depth$

06     $then \ depth \leftarrow a.depth$

07    $end \ if$

08    $if \ depth \geq 2N$

09     $then \ break \ while$

10    $end \ if$

11    $for \ \ t{:}T_0$

12     $if \quad a \in \Pr e(t) \quad and \quad \Pr e(t) \subseteq A_0 \cup A_d$

13      $then \ Flag \leftarrow false$

14       $for \ a^{'} : Post(t)$

15        $if \ \neg \exists \in A_0 \cup A_d , a^{'} \leq b$

16         $then \ a^{'}.depth = depth + 2$

17          $A_d \leftarrow A_d \cup a^{'}$

18          $waiting\_queue.push(a^{'})$

19          $E \leftarrow E \cup (t, a^{'})$

20          $Flag \leftarrow true$

21        $end \ if$

22       $end \ for$

23       $if \ Flag = ture$

24        $then \ t.depth = depth + 1$

25         $T \leftarrow T \cup t$

26         $E \leftarrow E \cup (\Pr e(t) \times t)$

27       $end \ if$

28     $end \ if$

29    $end \ for$

30    $depth \leftarrow depth + 2$

31 $end \ while$

Assume that there are n hosts in the network, the largest number of used vulnerability which exists in each host is c. Due to limit of the algorithm that each host can only be put into search queue once, the scan needs to match c

vulnerabilities and n hosts. According to the above analysis, we can conclude that the time complexity of algorithm is $O(cn^2)$, usually the number of used vulnerability can be ignored to the host number, so the time complexity of the algorithm can be expressed as $O(n^2)$ (n is the number of host).

4 Experiment and Analysis

To test the algorithm scalability, we design a contrast experiment in the large-scale network. Contrast with Sheyner and Ou algorithm, Ou algorithm use MulVAL inference engine to automatically construct logic attack graph, Sheyner algorithm based on model checking technology to achieve the state attack graph automatically generated. Attack graph generated by the two kinds of algorithm only contain atomic attack related to attack target, belongs to the part of the attack graph, and the attack graph by algorithm in this paper shows all possible attack results that the attacker with initial attack ability may adopt the attack path.

The experiment simulates five incremental large-scale networks to generate attack graph, firewall does not exist in the target network, all hosts can communicate with each other, and each host has three vulnerabilities at most. The number of host in five target network is 10, 20, 50, 100 and 200 respectively, CPU time that the algorithms generating attack graph in different scale target spend on is shown in Fig. 3. Line 1 indicates Sheyner algorithm, Line 2 indicates Ou algorithm, and Line 3 stands our method.
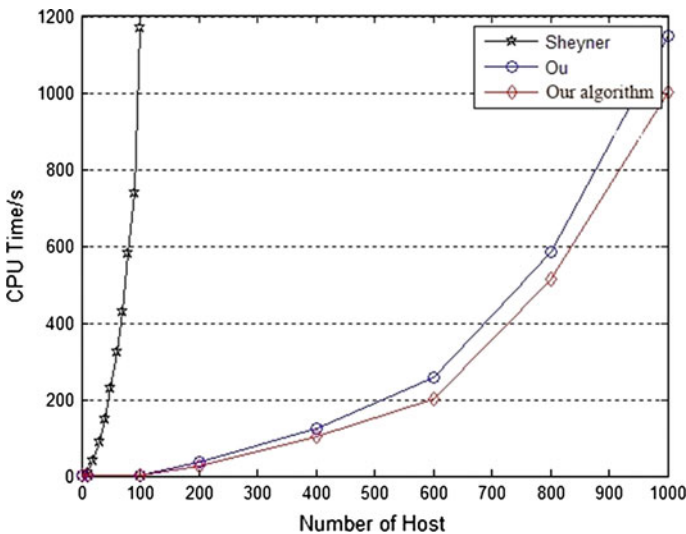

Fig. 3 CPU run time of each attack graph generating algorithm in different scale network

The experiment results show that actual performance of three algorithms in accordance with network scale grows exponentially, but the rate are different. Sheyner algorithm has poor scalability, and actual performance of our algorithm and Ou algorithm in accordance with network scale relatively grow slower, and they have better scalability. At the same time, the actual performance of our algorithm in this paper is better than that of Ou algorithm, because our algorithm limits the number of attack steps, which considers long path as invalid path, consequently optimize the algorithm performance.

References

1. Noel S, Jajodia S (2008) Optimal IDS sensor placement and alert prioritization using attack graphs. J Network Syst Manage 3(16):259–275
2. Lippmann RP, Ingols KW (2005) An annotated review of past papers on attack graphs. Technical Report, ESC-TR-2005-054, MIT Lincoln Laboratory
3. Jha S, Sheyner O, Wing J (2002) Two formal analyses of attack graphs. In: The l5th IEEE computer security foundations workshop. IEEE Computer Society, Cape Breton, pp 49–63
4. Ammann P, Wijesekera D, Kaushik S (2002) Scalable, graph-based network vulnerability analysis. In: The 9th ACM conference on computer and communications security. ACM Press, New York, pp 217–224
5. Wang L, Noel S, Jajodia S (2006) Minimum-cost network hardening using attack graphs. Comput Commun 29(18):812–824
6. Sheyner OM (2004) Scenario graphs and attack graphs. Carnegie Mellon University
7. Ou X, Boyer WF, McQueen MA (2006) A scalable approach to attack graph generation. In: Proceedings of the 13th ACM conference on computer and communications security, pp 336–345
8. Ou X (2005) A logic-programming approach to network security analysis. Princeton University, Princeton
9. Noel S, Robertson E, Jajodia S (2004) Correlating intrusion events and building attack scenarios through attack graph distances. In: Proceedings of the 20th annual computer security applications conference, vol 12. Tucson, AZ, USA, pp 350–359
10. Swiler LP, Philips C, Gaylor T (1988) A graph-based network-vulnerability analysis system. Technical Report. SANDIA Report No. SAND 97-3010/1
11. Qin X, Lee W (2004) Attack plan recognition and prediction using causal networks. In: Proceedings of international conference on computer security applications, Atlanta, USA, pp 370–379

Design of the Framework for Reverse Model Based on TTCN-3 Test Systems

Yongpo Liu, Ji Wu, Chuangye Chang and Shuangmei Liu

Abstract Aimed at the comprehensibility, reusability and maintainability, the thesis presents the reverse model recovery for the legacy code developed by TTCN-3. It can also help tester and maintainers to verify the test implement, etc. First, the thesis introduces reverse model and its features based on TTCN-3 test systems. Then, the thesis builds the framework on Eclipse platform using the plug-in mechanism. Here the thesis reuses and expands the core parser in TRex project for the reverse engineering analyzer.

Keywords TTCN-3 · Reverse engineering · Test system · Static analysis · Meta-model

1 Introduction

TTCN-3 is a test specification and test standard developed and promoted by ETSI. It is a kind of test description specification with rich ability based on black-box, and can be applied to various forms of distributed systems. As TTCN-3 develops, it has been increasingly applied to all kinds of testing fields. Testing systems development based on TTCN-3 has similar characteristics with software development. But as the test system grows in size and the testers change, for the huge test systems, the management and maintenance have become more and more difficult [1]. Therefore, the reverse engineering based on TTCN-3 test systems can

Y. Liu (✉) · J. Wu · C. Chang
School of Computer Science and Engineering, Beihang University, Beijing 100191, China
e-mail: liuypo@sei.buaa.edu.cn

S. Liu
Science and Information College, Qingdao Agriculture University, Qingdao 266001, China
e-mail: lshuangm@163.com

help testers grasp the system design from higher levels, and can test the consistence between test design and test implementation, which is of great significance and important value for test system maintenance, expansion and evaluation.

This thesis designs the system framework of reverse model discovery. Based on Eclipse, it achieves the framework extension and multiplexing very well by using a plug-in mechanism. Then by extending the TTCN-3 static analyzer in TRex project, static basic information, static test configuration, test data and call relation can be extracted by designing the corresponding interface. At the same time, the test trajectory can be obtained from the test case running on TTworkbench from Testing Technologies, and then the dynamic test profile extraction can be achieved [2]. This tool shows the basic information and abstract design of a test system to testers and maintainers from static and dynamic aspects, and can be used to maintain and update the test system.

2 Introduction of Reverse Engineering

Reverse engineering is an important branch of software engineering, along with the improvement of the complexity of software and the increase of legacy systems, reverse engineering receives more and more people's attention, which has a broad space of development. Chikofsky and Cross defined the reverse engineering as an process to analyze the target system, it generally comprises the following two parts:

(1) Identifying system components and analyzing the dependency relationships between them.
(2) Building additional or higher abstract level of expression forms of the system.

According to the above definition, reverse engineering is essentially an engineering of intellectual recovery of knowledge discovery. For the invisibly artificial product, the first purpose of reverse engineering is mainly to reproduce the original design knowledge of something studied, the second one is to find the design knowledge expressed unclearly in the design and existed in the designers' mind [3].

From the definition of reverse engineering, software reverse engineering includes analyzing, abstracting and displaying systems, which aims at helping users understand the objective system.

In this thesis, the traditional reverse method is applied to TTCN-3 test systems. The information can be obtained from two aspects, namely static and dynamic. Then the algorithm related to the areas is developed. The abstract model describing the test systems will be extracted and displayed. Reverse engineering helps testers and maintainers understand and maintain the test systems.

3 Reverse Engineering Based on TTCN-3 Test Systems

With the increasingly prominent role in software test, testing gets more attention as an important procedure in the software process. The traditional software test focuses on the manual test stage, which is more but duplication. As the automatic test technology becomes riper day by day, the software test has entered a stage of rapid development. But there isn't a unified standard on test description. At present, XML and TCL are widely used. It becomes a hot topic about how to make test more efficient, standardized and reusable [4]. When TTCN-3 was put forward by ETSI in 2001, this problem has been solved. The core language of TTCN-3 test specification is similar to traditional programming language, which can be used in wider fields to describe the test. It provides greater convenience and flexibility in writing test sets. Because TTCN-3 is driven by a test, its grammar is similar to traditional programming language, with a special test propagation characteristic. So compared with traditional programming language, TTCN-3 is more concerned with the following aspects, that is, testing judgment processing, the template matching mechanism between SUT motivation and the feedback information received expectedly, timer processing, test execution control mechanism, dynamic test configuration, synchronous or asynchronous communication, test process distribution mode and information coding ability, etc.

Since TTCN-3 language has powerful functions, such as dynamic configuration, synchronous and asynchronous communication, etc. testers find it difficult in ensuring the consistency between the test design and the test implementation; and it also needs to have the auxiliary tools when facing the management and maintenance of the huge test systems. These problems will be paid more attention with the mature of TTCN-3 test language. Reverse analysis of the test systems based on TTCN-3 can help testers and maintainers control the system design from higher level, and test the consistency between the test design and the test implementation, which has important significance for test system maintenance, expansion and assessment [5].

In essence, the reverse engineering based on TTCN-3 test system is a process to discover test design and test model. The core issue of reverse engineering is to model the target system, therefore how to build the meta-model for TTCN-3 system must be solved, on which the test system model will be generated. Then the algorithm related with some fields is applied to the test system model, which can abstract and extract the test design and test model.

4 Design the Framework of Model Discovery System

In general, reverse engineering includes the following two parts. The first is to identify and analyze the dependency relationship between components; the second is to establish the expression way with additional forms or at a higher level of

abstraction. Reverse engineering is essentially a process of intellectual recovery and knowledge discovery. So in this paper it is divided into three steps:

(1) Data extraction: This data is untreated, and is the metadata to get source file through static or dynamic analysis which is the basis of reverse analysis and model discovery.
(2) Knowledge organization: The extracted data is classified and saved. Here the metadata is restructured by using the already defined generic model or specifications, which can realize data construction.
(3) Information display: On the basis of obtained knowledge, the algorithm related to some fields is applied to the system, which can realize the information construction and extraction from higher levels.

To summarize the reverse engineering and model discovery techniques, in this paper the system framework of model discovery is designed for the reverse engineering based on TTCN-3 test language, as shown in Fig. 1.

The framework is based on Eclipse, it is easy to be maintained and extended by using the plug-in extension mechanism by Eclipse. At the same time the framework reuses a large number of infrastructures provided by Eclipse, which is in favor of system integration and development. The framework is divided into three parts: static analysis, modeling and displaying. The input of static analysis is the test set of TTCN-3, the output is the abstract syntax tree and the symbol table, also includes the dynamic trajectory extraction, namely obtaining the metadata from the test set; The input of modeling is U2TP and TTCN-3 specifications, in the process the meta-model of TTCN-3 test system is designed, the mapping rule from U2TP to TTCN-3 is implemented. The output is the TTCN-3 test system model,
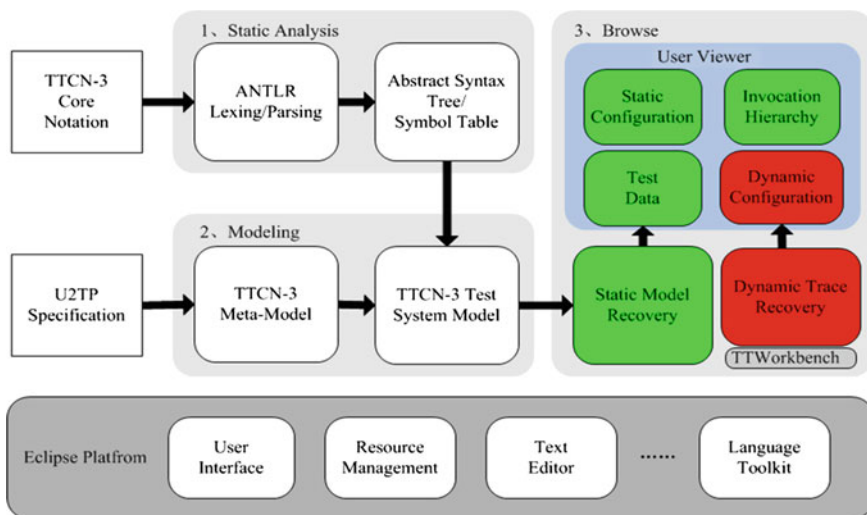

Fig. 1 The discovery framework of reverse model

namely realizing the data reorganization and the formation of information; the input of displaying is the TTCN-3 test system model, the test design and model discovery are implemented by applying a field-related algorithm. The output is various views of TTCN-3 test system, namely displaying the model.

5 Design the Meta-Model of TTCN-3 Test System

Although TTCN-3 provides such the table and graphical description form as TFT and GFT, it is only a format describing test cases, and can't show the overall structure and logic of a test system. In order to solve the problems about maintaining, improving and reusing the codes from a large number of legacy systems, it needs to provide a description method with higher level for TTCN-3. So, for the target of model discovery, this paper realizes the mapping relation between U2TP and TTCN-3, and designs the meta-model of TTCN-3 test systems by means of the specific concepts in U2TP [6].

Referring to the description of the test system in U2TP and the mapping rule defined above, the TTCN-3 test set is divided into four parts: A TestArchitecture, TestBehavior, TestData, Time. The meta-model of the test system is shown as Fig. 2.

In the design process, this paper considers the special demand of reverse engineering, and properly cuts the meta-model of TTCN-3, which is convenient for representing and displaying test system models. Simultaneously in order to facilitate the basic information organization and storage in the TTCN-3 test system, the related data structure used in recording test information is also defined in the meta-model. From Fig. 2, the root node of the whole system is TestSystem. Every test system has its own name, namely system, which is generally the name
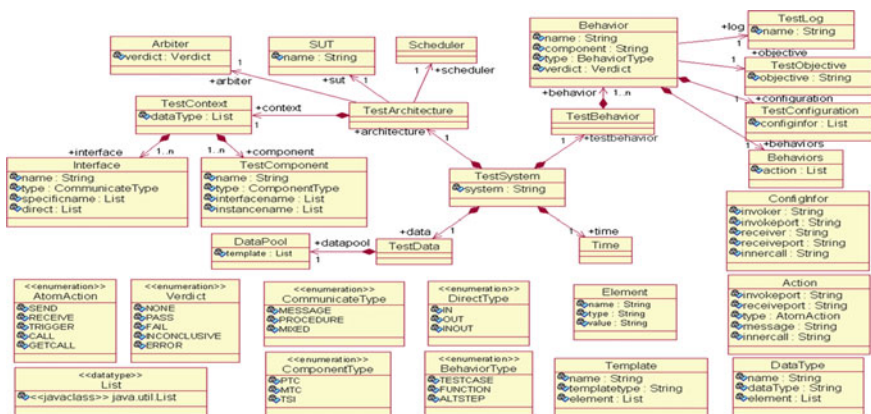

Fig. 2 The meta-model of TTCN-3 test system

of the abstract test set. A TestSystem logically includes four parts: TestArchitecture, TestBehavior, TestData and Time.

In addition, the meta-model of the test system also includes the specific characteristics of the test language, which are mainly the definition of some enumeration, including the atomic operation of communication AtomAction, the result of the judgment Verdict, the type of test communication CommunicateType, the type of test components ComponentType, the type of testing the communication direction DirectType, the type of test function BehaviorType, etc.

6 Design and Implement the Static TTCN-3 Analyzer

6.1 Compare Three Static Analyzers

In order to facilitate customizing the required reverse analysis tool, we mainly research the open source tool of the static TTCN-3 analyzer. The following is the comparative analysis of three static analyzers.

- ttthreeparser: It was an early analyzer developed by Testing Technologies in Germany, and currently only supports TTCN-3v1.1.2, which is one of the earliest version of TTCN-3. The current test scripts are based on TTCN-3v3.1.1, its grammar has been undertaken bigger adjustment, so it was of little value for this paper.
- ttcn3parser: It is a static analyzer based on Python, which is provided by Debian that is an open source organization. It supports TTCN-3v3.1.1. But due to the lack of appropriate interface and documents, the grammar needs to be rewritten, the workload is too much.
- TRex: It is developed by the Motorola and Germany's Gottingen university cooperation, and used to asset and reconstruct the TTCN-3 test sets, including the static analysis of TTCN-3v3.1.1 and the clear definition of interfaces and data structures. The basic information and model definition can be extracted through the static reverse analysis. TRex supports the existing standard of TTCN-3 very well, and is also an open source research project. In addition, TRex exists as Eclipse plug-in, which is very favorable for system expansion and inheritance.

6.2 Introduce TRex

TRex is a TTCN-3 metric and refactoring tool based on Eclipse plug-in mechanism. Positive TTCN-3 tools generally focus on the development of the core compiler and test actuator, TRex belongs to reverse engineering support tools, its goal is to analyze and optimize the test system codes. The emergence of TRex makes the

automatic measurement and reconstruction of TTCN-3 test codes become reality, it also provides favorable basis to evaluate the test system design [7]. However, TRex can't help testers grasp the test system design from higher level, and test the consistence between test design and test implementation. With the development of TTCN-3, more legacy systems need to be maintained and upgraded. How do the testers understand and grasp the existing test system in the shortest possible time, and test the consistence between the test design and the test implementation, it will become a major concerned issue. The research in this thesis meets the requirement of TTCN-3 development very well. Combined with TRex, it will provided strong support in maintaining and managing the TTCN-3 test systems.

TRex is mainly used for the measurement and reconstruction of the TTCN-3 test systems, and is divided into three modules, shown as Fig. 3.

- Input the TTCN-3 sets, the abstract syntax and symbol table can be obtained through static analysis, namely the metadata of the test sets.
- Input the metadata of TTCN-3 test sets, the metrics of the test sets can be obtained, which is the basis to evaluate the test sets.
- Input the metadata and metrics of the test sets, the test codes can be restructured by changing the syntax tree.

The static analysis, measurement and reconstruction provide the basis for maintaining the test sets, while the higher quality of TTCN-3 test sets can be output.

6.3 Extend TRex

Firstly the compiler transfers the source programs written by programmers into a convenient data structure, which is the process of lexical and syntax analysis. The
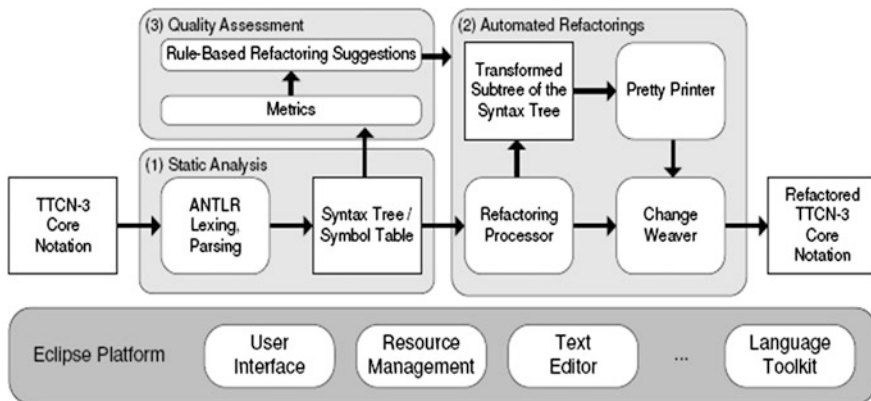

Fig. 3 Architecture of TRex

abstract syntax tree and symbol table can be obtained after TRex analyzes TTCN-3 source codes. The abstract syntax tree uses Tree as the data structure, because Tree has strong recursive after the any node of Tree is extracted. Node is still a complete Tree, which conforms to the formal language in the compiler theory, such as using function in the function, loops in the loop, conditions in the condition, etc. It can be easily expressed in Tree. Aiming at the problems to be solved, this paper mainly inspects the de.ugoe.cs.swe.trex.core.analyzer.rfparser of de.ugoe.cs.swe.-trex.core, which contains the TTCN-3 static analyzer. According ANTLR grammar analyzer by use of the toolkit provided by ANTLR. The abstract syntax tree of TTCN-3 can be obtained by analyzing the TTCN-3 source files by means of the static analyzer, then the information will be restructured by the use of symbol table.

The important data structures in TRex are Symbol, LocationAST, Scope and SymbolTable, shown in Fig. 4.

TTCN-3 defines rich data types, which make it convenient to define the test configuration, test behaviors and test data for users. TRex defines a Symbol for every TTCN-3 element, shown in Fig. 5.
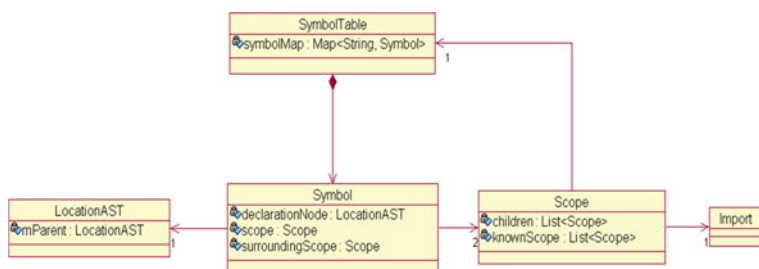
Fig. 4 Symbol table of TRex
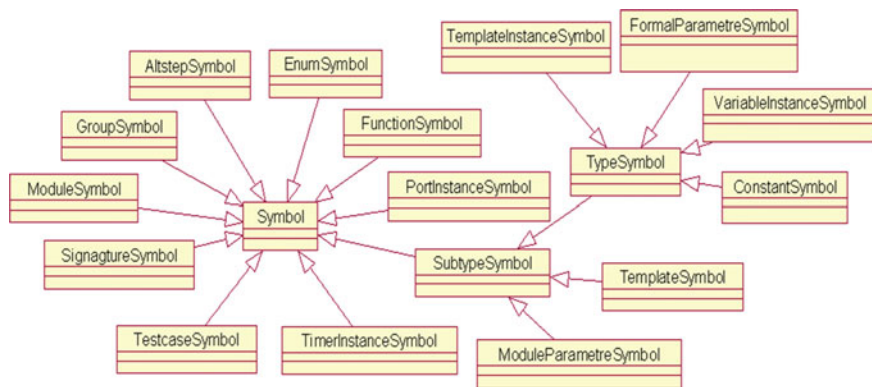
Fig. 5 Symbols of TRex

In the process of reverse analysis, this paper mainly analyzes the following key structures: ModuleSymbol, TypeSymbol, SignatureSymbol, EnumSymbol, PortInstanceSymbol, SubtypeSymbol, TemplateSymbol, TestcaseSymbol, FunctionSymbol, AltstepSymbol, etc.

TRex is mainly used for TTCN-3 test set metrics and restructure, the editing and static analysis function of the TTCN-3 scripts are the core of the tool, this paper also analyzes the TRex editor, the incident response policy model over the editor, the factory pattern of the static analyzer and entrance program of the TTCN-3 static analyzer. Based on the study of the core codes of the static analyzer, the data structure and interfaces defined by ourselves are added into TTCN2Analyzer, which is used to access the various information of the static analyzer, and then realizes the reverse discovery of the model.

7 Conclusions

This paper presents the design and implementation of the system framework of the reverse model discovery, and introduces the design and implementation of the static analyzer in detail. The reverse engineering on TTCN-3 test systems can help testers grasp the test system design from higher level, and can test the consistence between the test design and the test implementation, which is of great significance and important value for testing system maintenance, expansion and evaluation.

References

1. Schieferdecker I, Din G (2010) A meta-model for TTCN-3. In: 1st international workshop on integration of testing methodologies, ITM 2010, Toledo, Spain
2. Sun C, Zhou J, Cao J, Jin M, Liu C, Shen Y (2009) ReArchJBs: a tool for automated software architecture recovery of JavaBeans-based applications. In: Proceedings of ASWEC 2009, IEEE Computer Society, Brisbane, Australia, March 29–April 1, pp 270–280
3. Grabowski J (2011) TTCN-3-A new test specification language for black-box testing of distributed systems. In: Proceedings of the 17th international conference and exposition on testing computer software (TCS 2011), theme: testing technology vs. testers' requirements, Washington D.C., pp 231–240
4. Dai ZR (2004) Model-driven testing with UML 2.0. Second European workshop on model driven architecture (MDA) with an emphasis on methodologies and transformations (EWMDA'04), Canterbury, England
5. Peng S, Zhu Q (2006) Inverse modeling based on source code analysis [J]. Appl Res Comput 3:52–54
6. TRex WebSite [OL] (2006) http://www.trex.informatik.uni-goettingen.de
7. Baker P, Evans D et al (2010) TRex—the refactoring and metrics tool for TTCN-3 test specifications, TAIC-PART. In: Proceedings of the testing: academic and industrial conference on practice and research techniques, pp 90–94

The Information Acquisition System of Subsidence Mining-Induced Based on Mobile GIS

Meiwei Zhang, Weicai Lv and Guanghu Yao

Abstract In order to solve the low efficiency, low inefficiency and low informationization problem of current mining subsidence acquisition technology backwardness, this paper provides a reliable and practical data acquisition terminal for mining subsidence monitoring and prediction work under CORS space information framework. This paper also adopts application of network RTK on the PDA communication between terminal and data center communication, Mobile GIS data organization, terminal and data center data fusion and other contents. The practice shows that this System greatly improves the speed and accuracy of the monitoring points' information.

Keywords Mining subsidence · Network RTK · Date fusion technology · Movement and deformation · Mobile GIS

1 Introduction

Due to complex of mine terrain, big density of collection, high precision, the information of terminal get so hugeness. How to gain coal mine subsidence information fleetly and accurately has become a very important technical bottleneck. Mobile Geographic Information System (Mobile GIS) realize monitoring point information fleetly entering, editing, altering, drawing integration by electric,

M. Zhang (✉) · W. Lv · G. Yao
School of Geomatics, Anhui University of Science and Technology, Huainan 232001, China
e-mail: zhangmeiwei1@126.com

W. Lv
e-mail: 564271409@qq.com

G. Yao
e-mail: yaoguanghu@163.com

especially for space information, attribute information of monitoring point. This system combines CORS technology, using powerful GIS function. The data between it and platform of monitoring and forecasting can be mixed together and exchanged long-distance. Mobile GIS solves mining subsidence information collection technology problem effectively, guaranteeing mining subsidence monitoring and forecasting data accuracy, reliability and timeliness.

2 Mobile GIS Overall Design and Function Introduction [1–2]

According to the reality of coal subsidence and practical applicability of system, we choose Mobile GIS of ArcGIS Server 10, Visual Studio 2008 as development tool, C# [3] as development language to develop shift basing on .NET 3.0 environment, which technology roadmap is shown in Fig. 1. The system is mainly composed of nine modules below.

2.1 File Management Modules

File management including engineering file management and data collection management. Main function of File management: using text mode, database mode and graphics mode for memory management on different data mining project.

2.2 Data Acquisition Module

The data acquisition module mainly consists of attribute information collection, spatial information collection and topological information collection. In the field data acquisition, first, users entry point related attributes information through the graphics operation or GPS automatic navigation to the observation point; Then the measuring instrument equipment is connected to send equipment related instruction. The instruction will control equipment to collect data automatically and transfer the measurement data to the terminal; relevant data will be saved to the database or transferred remotely to coal mining subsidence monitoring and prediction system data center.

2.3 Data Management and Analysis Module

Data management and analysis module including the data processing and organization. Terminal data processing calculates summary deformation for different
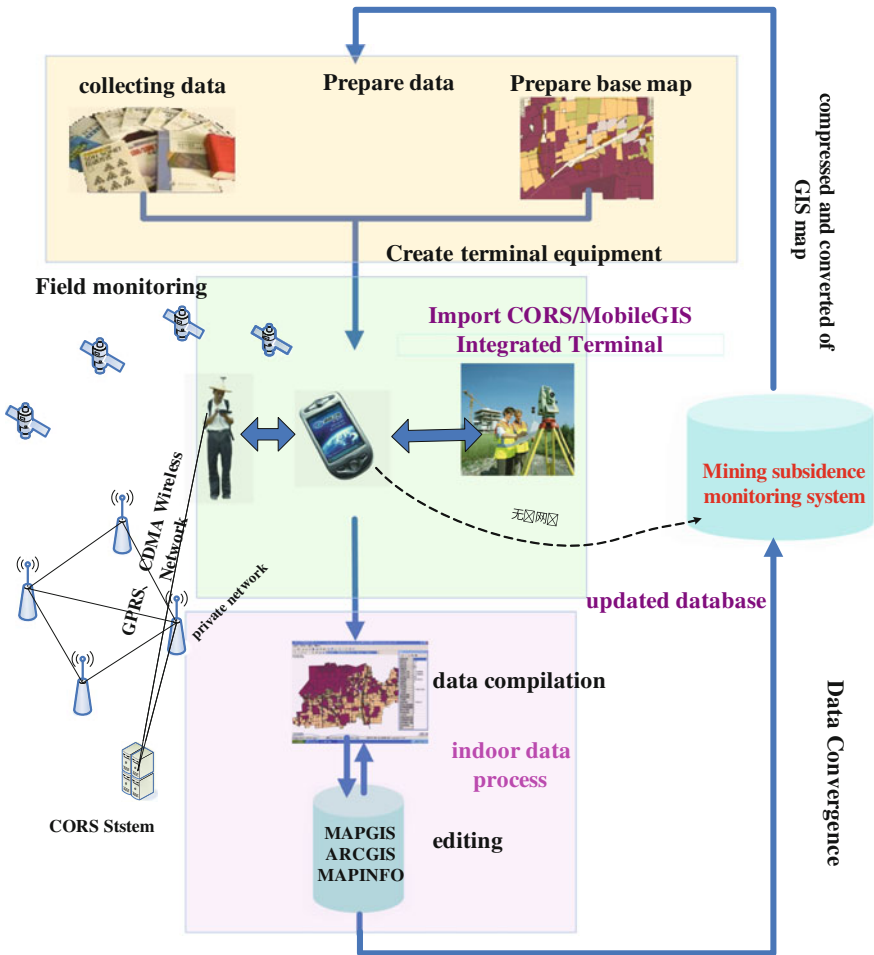
Fig. 1 Integrated technology road of system

periods of point position and surface movement deformation parameters using of all or part of the point. Data organization is to decentralized management which data collected on PDA in order to sort management different data. Each point on PDA can be set layer properties in which users can set different colors or control each layer.

2.4 Mobile GIS Module

Mobile GIS module is mainly responsible for loading and displaying base map, managing information of points, realizing operating base map.

Managing information is mainly to edit points' attribute information and convert geographic coordinates to screen coordinates and display them when plotting and drawing topological relation among points.

Operating base map including: Graphic amplification, reducing, translation and full figure operation. Mobile GIS provides a very important control Map, which operated in setting Map's MapAction attribute. MapAction consist of Zoom-InMapAction, ZoomOutMapAction, ZoomInOutMapAction, PanMapAction and other members. Users can navigate to the current collection area through simply operating base map.

2.5 Query Editor Module

Query editor module is an important means of the interaction of users with database information. Query editor module include: information query and edit the monitoring points. Users can edit attribute information when entry monitoring points, such as the monitoring point number, name, and installment number. If users need to view or modify the information of a monitoring point, it can be queried the point by providing a number of monitoring points or specifying a point on the diagram, and users can modify this information and save it to the database.

2.6 Measuring Instrument Integration Module

The system support measurement methods of total station and Network RTK based on CORS in order to meet users, convenient users fitting actual measurement. PDA and measuring instrument communicating by com communication which connection type is provide data link and Bluetooth connection through Bluetooth virtual com port communication.

2.7 Network Communication Module

If the Information Acquisition System of Subsidence Mining wants to achieve settlement of rapid transmission of information, it must be achieved through the network.

The GPRS functionality provided by the SIM card can be used in terminal acquisition system on the PDA [4]. The PDA can establish a connection with the data center by accessing to the Internet network through the GPRS network. It can be connected to the Internet directly via a wired in the data center to realize the sharing of the mobile end data center data.

2.8 Coordinate Conversion Module

Before the monitoring point measurement, we need to set the coordinate system. GPS and network RTK positioning results are based on the WGS84 coordinate system [5]. In order to meet the needs of the survey work, the system must establish a conversion model between the WGS84 coordinate system with 54 countries 80 countries in 2000 and the local coordinate system.

2.9 Navigation Module

Navigation module [6] is designed for the users to navigate to the location of the current user location corresponding graph quickly and conveniently. The principle of the navigation module implements is as follows: It receive GPS signal by GPS serial port in PDA, and extract rough coordinates from the GGA statement which is efferented by GPS module built-in the PDA. Then the coordinates will be converted to the matching three-dimensional coordinate by the dynamic coordinate conversion module in the real-time, and the map will be enlarged the appropriate ratio and displayed at the center of the MAP control in the area of three-dimensional coordinates by calling the Mobile GIS module.

3 Ntrip Protocol

NtripClient communicate with CORS Center by GPRS Network, it's working principle as follow.

3.1 Obtain Resources List

NtripClient send request message to CORS center, if NtripSource/mountpoint is not exist, NtripCaster will response: "SOURCETABLE 200 OK", then send "source-table": e": <Source-Table> .NtripClients could select NtripSources/ mountpoints from all available NtripSources/mountpoints.

The message is:

```
string msg = "GET/HTTP/1.1\r\n";
msg += "User-Agent: " + <server >+"\r\n";
msg += "Accept: */*\r\n";
msg += "Connection: close\r\n";
msg += "\r\n";
```

3.2 User Authentication

Users select mountpoint that they need, then send the second message to CORS center. If the mountpoint is exist and username and password is correct, CORS center will send GNSS datas to NtripClient.

In order to get authorization, client must send user ID and password with base64 code. If client want to send "Aladdin" and "Open sesame", he will use the header bellow:

→Authorization: Basic QWxhZGRpbjpvcGVuIhNIc2FtZQ==

The second message is:

```
string msg = "GET/" +<mountpoint>+ " HTTP/1.1\r\n";
msg += "User-Agent: "+ <server >+"\r\n";
msg += "Accept: */*\r\n";
msg += "Connection: close\r\n";
if (_username !=null && _username != "")
{
   string auth = ToBase64(<username>+ ":" +<password>);
   msg += "Authorization: Basic" + auth + "\r\n";

}
msg += "\r\n";
```

3.3 Send Preliminary Position

Some Netword need send location information of NtripClient to NtripCaster.

NtripCaster adds the location information to the data stream and send to VRS. Ntrip allows client to request data stream and to send GNSS data on HTTP. If the parameter of <nmea> is "1", CORS center will send GNSS data after it gets NMEA GGA statement.

4 Data Processing and Analysis

4.1 Calculation of Movement and Deformation

Terminal data processing including the following movement and deformation calculation [7]:

4.1.1 Subsidence

$$W_n = H_{n_0} - H_{n_m} \tag{1}$$

where W_n is the subsidence value of point n and H_{n_0}, H_{n_m} respectively indicate the first time and the mth time elevation of point n.

4.1.2 Horizontal Movement

$$U_n = L_{n_m} - L_{n_0} \tag{2}$$

where U_n is horizontal distance of point n, L_{n_0} and L_{n_m} respectively indicate the first time and the mth time distance between control point and point n.

4.1.3 Inclination Deformation

$$i_{n-(n+1)} = \frac{(W_{n+1} - W_n)}{l_{n-(n+1)}} = \frac{\Delta W_{(n+1)-n}}{l_{n-(n+1)}} \tag{3}$$

where W_n, W_{n+1} are subsidence value and $l_{n-(n+1)}$ is horizontal distance of points n and n+1.

4.1.4 Curvature Deformation

$$K_{(n-1)-n-(n+1)} = \frac{\left(i_{n-(n+1)} - i_{(n-1)-n}\right)}{0.5\left(l_{(n-1)-n} + l_{n-(n+1)}\right)} = \frac{2\Delta i_{(n-1)-n-(n+1)}}{l_{(n-1)-n} + l_{n-(n+1)}} \tag{4}$$

where $i_{(n-1)-n}$, $i_{n-(n+1)}$ are inter-site average gradient and $l_{(n-1)-n}$, $l_{n-(n+1)}$ are inter-site horizontal distance of points n−1, n and n, n+1.

4.1.5 Horizontal Deformation

$$\varepsilon_{n-(n+1)} = \frac{(U_{n+1} - U_n)}{l_{n-(n+1)}} = \frac{\Delta U_{n-(n+1)}}{l_{n-(n+1)}} \tag{5}$$

Table 1 The trend point of observation movement and deformation result

Point number	Vertical movement and deformation			Horizontal movement and deformation	
	Subsidence (mm)	Inclination (mm/m)	Curvature (mm/m^2)	Horizontal movement (mm)	Horizontal deformation (mm/m)
CL06	+16.7	+0.37		78.7	+0.02
CL04	+25.0	+0.74	+0.008	+79.1	+1.23
ML69	+74.9	+2.69	+0.029	+162.2	+2.22
ML68	+255.7		+0.045	+311.3	
CL03	+0.3	-0.02		+9.5	-0.02
CL02	-0.5	-0.01	0.000	+18.6	+0.11
CL01	0.0	0.00	0.000	+10.0	-0.09
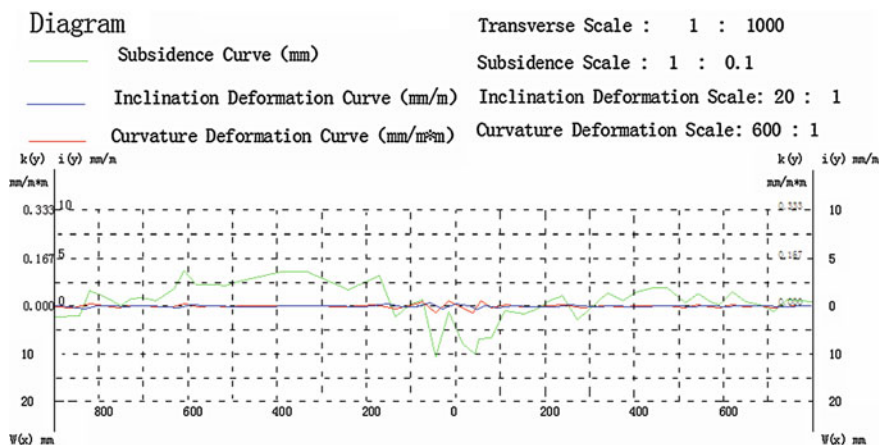
Fig. 2 Vertical movement and deformation curves of Zhuji coal mine

4.2 Result Output

Take a certain period data of Zhuji coal mine as the example, from (1)–(5) formulaes we obtain schema deformation values of the trend point of observation (Table 1).

We can draw Vertical movement and deformation curves below by the calculated deformation value (Fig. 2).

5 Summary

The Information Acquisition System of Subsidence Mining is an important part of Data acquisition system of Huainan coal mining subsidence monitoring system based on GIS technology. It can realize the mining subsidence monitoring data

on-site rapid acquisition, pretreatment, deformation analysis and remote transmission and provide timely and accurate monitoring data for the data center of mining subsidence monitoring and prediction system, it also realized the dynamic monitoring and management of regional mining area surface subsidence to provide technical support for environmental management of mine geology of the mining and government departments.

References

1. Zhao XC (2011) Research of data collection system based on mobile GIS and GPRS. School of Resources and Environment Science in East China Normal University, Shanghai
2. Zhong FG, Li N, Chen ZH (2010) Development and design of mobile GIS based on embedded language. Geomatics Spat Inf Technol 33:92–95
3. Nagel C, Evgen B (2010) C# advanced programming, 7th edn. Tsinghua University press, Beijing, pp 1334–1362
4. Li H (2011) PDA-based data processing system for construction site. School of Computer Science and Engineering Anhui University of Science and Technology, Anhui
5. Lv WC, Zhang RH et al. (2012) Coordinate system transformation for deformation monitoring datum network based on robust estimation. J Hefei Univ Technol (Nat Sci) 02
6. Dang YM, Mi JZ et al (2007) The principle and application of global navigation satellite system. Publishing House of Surveying and mapping, Beijing
7. He GQ, Yang L (1991) Mining subsidence. China University of Mining and Technology press, Xuzhou

Edge Collapse Considering Triangular Mesh for Model Simplification

Ruifang Zhu and Wei Shen

Abstract Because of re-triangularization during the process of traditional edge collapse, this paper introduces a new edge collapsing method improved with the use of triangular mesh. We deal with a series of feature points, and directly establish the triangular mesh which meet the need of simplification, and avoid triangulating secondly. The experiment shows that this method can improve the convergence rate, and simplified effect, under condition of assurance of the triangle mesh approximation precision and the decrease of the amount of the triangular facets.

Keywords Edge collapse method · Triangular mesh · Feature points · Accuracy analysis

1 Introduction

The methods based on vertex decimation and edge collapse are generally using in mesh model simplification [1]. These algorithms revealed some defects in different with the deepening of application. The edge collapse is very similar to the vertex decimation, except for re-triangularizing the holes after simplification, as thus reduce the effect of model simplification. In order to avoid the above defect, there is a kind of improved edge collapse algorithms, namely triangle mesh method. The triangular facet is the deleted element in simplified process, refrain from

R. Zhu · W. Shen (✉)
College of Marine Science, Shanghai Ocean University, Shanghai 201306, China
e-mail: wshen@shou.edu.cn

R. Zhu
School of Communication and Information Engineering, Shanghai University,
Shanghai 200070, China

re-triangularization consequently. Experimental results illustrate that the edge collapse method considering triangular mesh can directly improve simplified efficiency and maintain features of model.

2 Relation Work

The vertexes are firstly divided into simple vertex, complex vertex, boundary vertex, and interior point and angle point located in the edge feature according to the local geometric topological characteristics about the traditional edge collapse [2]. Then edge collapse algorithms first according to the model local geometry,and then according to all kinds of the error between the original model and simplified one, decided to the vertex that would be deleted. If the accuracy is smaller than the error threshold, the vertex was deleted, meanwhile make the holes cause by the deleted vertexes triangularized (Fig. 1).

3 Improved Method

This paper proposed a new simplified algorithm of edge collapse integrated with triangular mesh. The main idea of this method is: adding the judgments to the feature edges and feature points among the mesh in each step of simplified operation of edge collapse, and maintain the feature point as decimating the edge, then all of the points linked with the deleted edge were connected with the candidate points, make the model still keep the features of triangular mesh.

3.1 Extracting the Edge Feature

Normally, edge weighted value is often used to extract the feature edges. Firstly, calculate the weighted value of every edge, and decide which feature edge by a threshold is. Here, two methods are introduced [3]:

(1) SOD(Second Order Different) (See Fig. 2).

Fig. 1 The principle of edge collapse algorithm
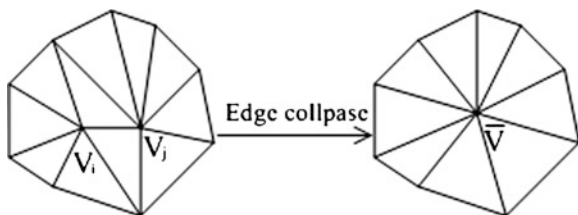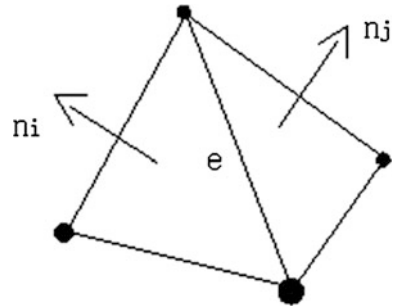
The method is simple, the dihedral angle of each non-border edge,

$$w(e) = \arccos\left(\frac{n_i}{\|n_i\|} \cdot \frac{n_j}{\|n_j\|}\right) \tag{1}$$

Here, n_i, n_j is the normal of two triangles linked to the edge e respectively. Given a angle threshold θ, when $w(e) > \theta$, e is a feature edge.

(2) ESOD(Extended Second Order Difference) (See Fig. 3).

ESOD is a method that consider the neighborhood of the edge e:

$$w(e) = \arccos\left(\frac{n_1}{\|n_1\|} \cdot \frac{n_2}{\|n_2\|}\right) \tag{2}$$

Here, n_i, n_j, is the normal of two terminal vertexes of the edge e respectively. Just like SOD, compare $w(e)$ and angle threshold θ, if $w(e)$ is greater than θ, e is feature edge.

Hubeli and Gross [4] analysis SOD and ESOD methods, and think that the SOD is easy and can get the important mesh characteristics when the model' size is not large. Meanwhile, SOD is efficient, however, when dealing with more complex mesh, such as geographic mesh model, SOD method will mistake noise as feature.
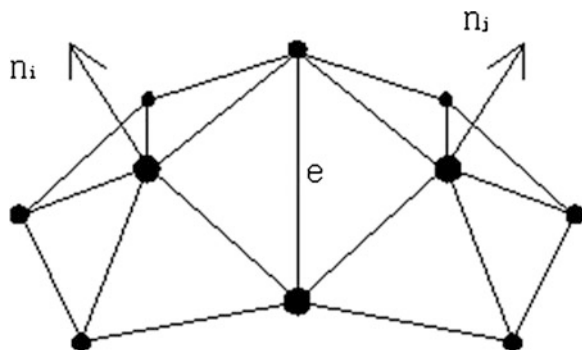
Fig. 4 The diagram of
vertex v
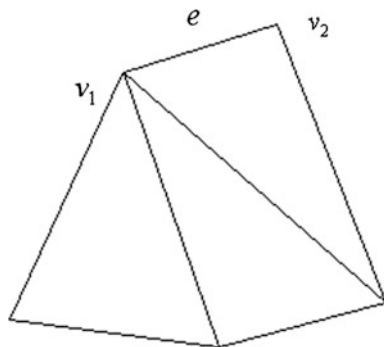
ESOD method compared with SOD, use more local information to determine the
feature edge, enhance the capability of noise resistance and improve the accuracy
of simplification.

3.2 Extracting the Feature Point

The method of angle deviation is used to extract feature point and maintain
characteristic of mesh, meanwhile it can judge the top point [5]. Defined the angle
deviation of vertex v as:

$$d(v) = b\pi - \sum \varphi_i \tag{3}$$

Here, φ_i is the angle of all triangles linked to v, and v is the vertex of those
angles. As shown in Fig. 5, when the vertex v is interior point, the value of b is
taken at 2, when the vertex v locate on the mesh boundary, b for 1. Extracted top
point through setting the threshold value of $d(v)$ (Fig. 4).

Fig. 5 The angle deviation
of vertex v
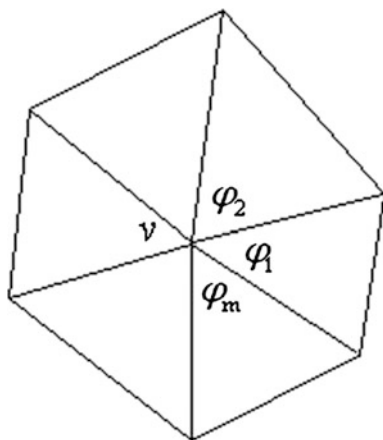
According to the vertex normal calculated by above mentioned ESOD methods, we will be able to extract the feature edges of the model, combined with simple threshold method to determine the type of mesh vertices, extracting the point of the features edges, finally get most of the feature points, and then find the top points as long as by using the angle deviation, just doing like that, we can complete the extracted feature points on the model. For references, follow the given guidelines.

4 Analysis of the Simplified Accuracy

As usually, the original model is sampled firstly, and then use the set of the sampling point to approximate the simplified error of calculated model, so here take 50 % of the set of any origin model. The error metric can be written as:

$$E_{\max}(M_1, M_2) = \max_{v \subset m} \ d_v(M_2) \tag{4}$$

$$E_{evg}(M_1, M_2) = \frac{\sum_{i=1}^{m} (d_v(M_2) \cdot s_i)}{\sum_{i=1}^{m} s_i} \tag{5}$$

Here, M_1 is the original model, M_2 is the simplified model, m is the set of sample points of M_1, $d_v(M_2)$ is the distance from the sampling point to the simplified model M_2, s_i is triangle area of sampling points on $M_1 \cdot E_{\max}(M_1, M_2)$ is the maximum error for the simplified model. $E_{evg}(M_1, M_2)$ is its average error.

5 Experiment Results

We apply our method to the Stanford university graphics laboratory Cow model (Fig. 6a). (a) is the original model shown in software MeshLab [6]. Firstly we use traditional edge collapse algorithm and our method to simplify Cow model. The number of triangle facets in the original model is 4762. (b),(c),(d) are the simplified model, whose simplification radio is 50, 70, 90 % respectively. The left columns of Fig. 6 are the results of traditional edge collapse algorithms, the right columns are results of our algorithm.

Compared the left column and right one of Fig. 6, we can see the mesh of cow head in the left column is uniform after simplified; features are not particularly obvious, while the right column cow eyes were simplified less, clearer than the left column. The cow nipple in the left column has been deleted when simplified rate of 90 %, and cow nipples still holds in the right columns.

Table 1 lists the maximum error of simplified cow model algorithm based on edge collapse and our algorithm; Table 2 is the average error.
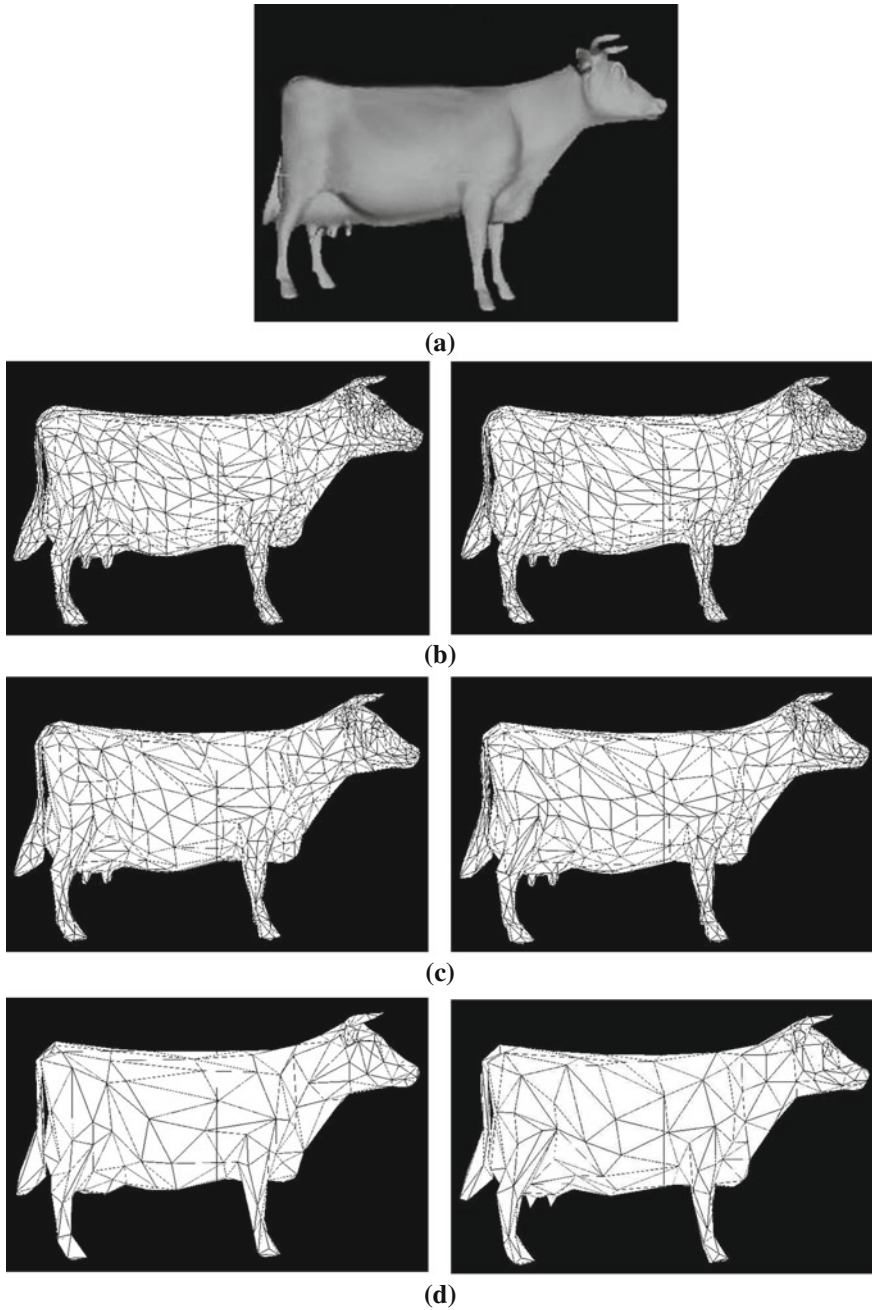
Fig. 6 The experimental results of simplified cow mesh. **a** The original mesh of cow. **b** The number of *triangular facets* is 2,908. **c** The number of *triangular facets* is 1,726. **d** The number of *triangular facets* is 620

Table 1 The maximum errors of simplified cow model (cm)

Ratio of simplification	50 %	70 %	90 %
Edge collapse	7.73	10.22	16.16
Our algorithm	7.71	8.83	12.50

Table 2 The average errors of simplified cow model (cm)

Ratio of simplification	50 %	70 %	90 %
Edge collapse	1.90	2.41	3.91
Our algorithm	1.89	2.30	3.55

The data of Tables 1 and 2 shows that the average errors of the simplified model got by our method generally slightly are smaller than edge collapse algorithm and the errors gap between two methods are larger with the simplified rate increasing. All this certifies that the accuracy of our method is in proportion to the simplified rate. It is thus clear that our method has a good practical and application.

6 Conclusion

This paper study the simplification method triangular mesh model, and discuss the mesh simplification based on edge collapse, and briefly describes the feature edges extraction, model features maintaining, calculate the accuracy of simplification, finally we verify our method by a test. Experimental results show that improved methods used in this paper are basically achieve the desired effect, keep the mesh stability, improved accuracy, can simplify the grid model quickly and meet the need of model accuracy.

References

1. Schroeder WJ, Zarge JA, Lorensen WE (1992) Decimation of triangle mesh. In: Proceedings of SIGGRAPH'92, pp 65–70
2. Hoppe H (1996) Progressive meshes. Computer graphics proceedings, annual conference series, ACM SIGGRAPH 1996, New Orleans, Louisiana, USA, pp 99–108
3. Liu X (2007) Finite element triangular mesh simplification. Dalian University of Technology, Dalian
4. Hubeli A, Gross M (2001) Multiresolution feature extraction from unstructured meshes. In: Proceedings of the IEEE visualization, Switzerland
5. Jiao XM, Michael T (2002) Feature detection for surface meshes. Numerical grid generation in computational field simulations
6. Meshlab. http://meshlab.sourceforge.net/

Research on Real-Time Security Risk Management of Rural Power Based on Monte Carlo Simulation Method

Xiaoqiang Song, Xia Lv, Xubo Guo and Zuhai Zheng

Abstract With the development of science and technology, reliability of power grids in rural areas in the algorithm have made new progress, in a variety of simulation methods, Monte Carlo simulation method because the sampling frequency has nothing to do with the grid size and the required accuracy only the features, so network security risk management in rural areas has been widely used. This article describes the sequential simulation, non-sequential Monte Carlo simulation of the basic principles of probabilistic simulation methods and characteristics.

Keywords Monte Carlo simulation · Power system · Risk assessment

1 Introduction

Rural power grid automation level unceasing enhancement with the development of economy, the deepening of the modernization,reliability of people on electric power is more and more strong,agricultural production and people's life will not normal without acceptable quality of power supply, so people demand on power

X. Song (✉) · Z. Zheng
College of Information and Electrical Engineering, Shen yang Agricultural University, Shenyang, Liaoning 110866, China
e-mail: 546020347@qq.com

X. Lv
Engineer, Master, Liaoning Province Agricultural Mechanization Institute, Shenyang, Liaoning, China

X. Guo
Shenyang, Liaoning Branch, China Mobile Communications Group Co., Ltd, Shenyang, Liaoning, China

reliability is higher and higher, and also prompted the electric power industry to find better ways to improve rural power grid reliability level.

Power system reliability research have a new progress in the reliability algorithm field, monte carlo simulation method which is a typical one, the application in power system reliability field which has maked improvement and enhancement in the new reliability algorithm. This paper will mainly discuss the characteristics of monte carlo simulation method [1].

2 Principle and Model

2.1 Non-Sequential Monte Carlo Simulation Method

The non-sequential monte carlo simulation method is often called the state sampling method,It is widely used in the risk assessment of power system. This method theoretical basis for the next system state is the combination of all components, and each element components state can be determined by sampling to the appear in this state probability.

Each element can be simulated by a uniform distribution in [0, 1] interval. Assuming that each element has two state (failure and work), and component failure is independent of each other, Make s_i representative state of element I, Q_i represents its failure probability, the first I element produce a random number R_i which in [0, 1] interval evenly distributed [2].

$$s_i = \begin{cases} 0, & \text{if } R_i > Q_i \\ 1, & \text{if } 0 \leq R_i \leq Q_i \end{cases} \qquad (1)$$

With N element of the system status by vector S presentation.

$$S = (s_1, \ldots, s_i, \ldots, s_N)$$

A system status was selected in sampling, the system analysis to judge whether it is failure state, if it is, then the state of risk index function estimation.

When sampling quantity is big enough, the system state S sampling frequency can be used as the probability of the unbiased estimation, such as formula (2) said

$$P(s) = \frac{m(s)}{M} \qquad (2)$$

Among them, M is sampling number; M (s) is system state s frequency in the sampling [3].

The non-sequential simulation sampling frequency speed faster, the algorithm is relatively simple, but because did not consider system sequence characteristics, so the calculation results are generally not include the reliability index that frequency of the fault occurrence and duration, Therefore, we joined the element

repair function into the non-sequential simulation, it is concluded that the computer frequency change interval method after a strict derivation. Because in the actual operation process using some simplified measures and sampling frequency limited, generally it is difficult to accurately calculate the exact value of system failure frequency index, so the Upper and lower limits of calculation frequency variation has important reference value.

2.2 Sequential Monte Carlo Simulation Method

Sequential monte carlo method is a simulation method which in a time span and according to sequence, Among them for establishing virtual system state transition cycle process have different methods, The most common is called state duration sampling method. state duration sampling method is based on sampling the element probability distribution of state duration, first specify all components of the initial state, usually assume that all element in running state, to sampling for in current status continuous time of each element. Should set probability distribution of condition duration for different state, such as operation or repair process, it can be assumed that has a different state duration probability distribution [4]. For example, the formula (3) is the sampling value of exponential distribution duration:

$$D_i = \frac{1}{\lambda_i} \ln R_i \qquad (3)$$

Among them, R_i is corresponding to the first I component in [0, 1] interval evenly distributed random number; if the current state is running state, λ_i is the first I component success rate; if the current state is decommissioning state, λ_i is the first I component failure rate; and if the current state is decommissioning state, λ_i is first I component repair rate.Secondly, in the research of time span (a large amount of sampling years) repeat, and record all components of each state duration sampling value, can be achieved the succession of the state transition process of every component in a given time span [5].

Combine all components state transition process, to establish system timing state transition cycle process. Because of using this method in simulation for many years, we can calculate the average reliability index, when flux index meet the requirements, the index will be designated reliability index for system. Sequential simulation can be used in the production simulation, and can be counted into different power failure time point, fault time length, the loss of cost by fault power failure caused. Similarly, also can be used to calculate some economic indicators.

Defect of sequential simulation that is needed a lot of calculation, with relatively low accuracy, but can increase the simulation time to improve the accuracy.

3 Monte Carlo Sampling Calculation Accuracy

For any element in the system I, Assumptions forced outage rate λi, Xi is its running state, so Xi is about the probability function P (Xi)

$$P(Xi) = \begin{cases} \lambda i & Xi = 0, \quad \text{element failure} \\ \\ 1 - \lambda i & Xi = 1, \quad \text{element normal operation} \end{cases} \tag{4}$$

If the system contain m components,$X = (X1, X2, \ldots, Xm)$ is a sample of system running state, $P(X)$ is a sample of the joint probability distribution function [6]. For every sample i Calculating reliability indicator function $F(Xi)$, $F(Xi)$ can be power outages mark(0 or 1)or Electricity power etc. In the theory, Ω state space all the sample after sampling,can use type (5) type (6) calculation of reliability index mean E (F) and variance V(F):

In the actual sampling, the sampling frequency restriction,get just $E(F)$ and V(F) of estimate $\widehat{E}(F)$ and $\widehat{V}(\widehat{E}(F))$:

$$\widehat{E}(F) = \frac{1}{N} \sum_{i=1}^{N} F(X_i) \tag{5}$$

$$\widehat{V}(\widehat{E}(F)) = \frac{1}{N-1} \sum_{i=1}^{N} (F(Xi) - \widehat{E}(F))^2 \tag{6}$$

N is actual sampling frequency.$\widehat{E}(F)$ is $E(F)$of Unbiased estimation,but $\widehat{V}(\widehat{E}(F))$ and $V(F)$ have the existence of the relationship [7]:

Commonly used at home and abroad and the variation of the sample ratio β as sampling process convergence criterion.

$$\beta = \frac{\sqrt{\widehat{V}(\widehat{E}(F))}}{\widehat{E}(F)} = \sqrt{\frac{V(F)}{N\widehat{E}^2(F)}} \tag{7}$$

when β is less than the given value, the suspension sampling. By type (11) type (12) still can have required sampling number [8].

$$N = \frac{V(F)}{\beta^2 \widehat{E}^2(F)} \tag{8}$$

By type (8) can be seen, Reliability index of the accurate estimation and the relevant system performance. System more safety,and $\widehat{E}(F)$ more small,The more difficult to carry on the estimate,The sampling frequency is more also. In a given convergence criterion, need the number of sampling and sample variance direct

Table 1 Reliability evaluation results

N	$L(10^{-3})$	$V(10^{-3})$	λ_L	EPNS	Ve	λ_E
5,000	9.10000	9.22647	0.157873	0.103000	2.02010	0.200101
10,000	9.79998	9.80401	0.100601	0.120001	2.42007	0.133245
15,000	9.87701	9.85301	0.082802	0.134778	2.69998	0.118793
20,000	9.46112	9.46801	0.074268	0.133562	2.76874	0.098875
25,000	9.51987	9.41938	0.070119	0.127300	2.56732	0.089645
30,000	9.50004	9.40865	0.059842	0.113580	2.40607	0.079905
35,000	9.40000	9.32546	0.056832	0.114065	2.29546	0.074258
40,000	9.14655	9.09293	0.052846	0.112860	2.18486	0.067846
45,000	9.25647	9.16787	0.049541	0.113984	2.48160	0.066254
50,000	9.27694	9.18583	0.046431	0.111936	2.58760	0.062453

proportion, reduce sample variance, can support the sampling frequency and enhance the computing efficiency.

4 Based on the Monte Carlo Simulation Algorithm Application

This method is to use the calculation accuracy with the sampling frequency increased change rule, system contains 12 units, 7 bus, 10 lines of small system, evaluate the results are shown in Table 1.

In the table, N for sampling frequency, V, Ve are respectively for reliability index L, EPNS to the test function of variance, λ L, λ E for L, EPNS corresponding variance coefficient. By the above data can be seen in the sample, test function of variance basic people quality constant, this is because the function of variance and experimental function characteristics and state variables (probability, the distribution function closely related, and the number of sampling and has nothing to do. With the increase of the number of sampling, the estimate of reliability index is more or less stable, variance coefficient gradually decreases, and improve calculation accuracy.

5 Summary

This paper studies the Monte Carlo exact solution approximation algorithm, and through the practical example is given to illustrate the application of Monte Carlo simulation method. Generally speaking in certain accuracy requirement, the Monte Carlo method sampling frequency and the size of the system has nothing to do, so is suitable for application in the rural power system reliability evaluation calculation. Based on the Monte Carlo method power system reliability assessment

process mathematical model is relatively simple and easy to simulation, so the result of calculation is more tally with the actual.

References

1. Kim TY, Oh KJ, Sohn I et al (2004) Usefulness of artificial neural networks for early warning system of economic crisis. Expert Syst Appl 264, 26(4):583–590
2. Jennings NR, Sycara K, Wooldrige M (1998) A roadmap of agent research and development. J Auton Agent Multi-Agent Syst 1, 1:275–306
3. Swaminathan JM, Smith SF, Sadeh NM (1998) Modeling supply chain dynamics: amultiagent approach. Decis Sci 29(3):607–632
4. Bauer B, Muller JP, Odell J (2000) An extension of UML by protocols for multi-agent interaction. International conference on multi-agent systems (ICMAa'00), pp 207–214
5. Huaget M-P (2002) Extending agent UML protocol diagrams. Technical report UL CS-02-009, vol 8. Department of Computer Science, University of Livepool, pp 35–37
6. Chiang HD, Flueck A, Shah KS, Balu N (1995) CPFLOW: a practical tool for tracing power system steady-state stationary behavior due to load and generation variations. IEEE Trans Power Syst 10(2):623–634
7. Ejebe GC, Tong J, Waight JG, Frame JG, Wang X, Tinney WF (1998) Available transfer capability calculations. IEEE Trans Power Syst 13(4):1521–1527
8. Gravener MH, Nwankpa C, Tai-Sim Y (1999) ATC computational issues [electricity supply]. Systems sciences, HICSS-32. In: Proceedings of the 32nd annual Hawaii international conference on power systems, vol 6

Canonical Duality for Radial Basis Neural Networks

Vittorio Latorre and David Yang Gao

Abstract Radial Basis Function Neural Networks (RBF NN) are a tool largely used for regression problems. The principal drawback of this kind of predictive tool is that the optimization problem solved to train the network can be non-convex. On the other hand Canonical Duality Theory offers a powerful procedure to reformulate general non-convex problems in dual forms so that it is possible to find optimal solutions and to get deep insights into the nature of the challenging problems. By combining the canonical duality theory with the RBF NN, this paper presents a potentially useful method for solving challenging problems in real-world applications.

Keywords Radial basis functions · Neural network · Canonical duality

1 Introduction

The problem of regression consists in trying to approximate a function $f : \mathbb{R}^n \to \mathbb{R}$ by means of an approximation function $g(\cdot)$ that uses a set of samples defined as:

$$\mathcal{P} = \{(x^p, y^p), x^p \in \mathbb{R}^n, y^p \in \mathbb{R}, p = 1, \ldots, P\}, \tag{1}$$

where (x^p, y^p) are respectively arguments and values of the given function $f(\mathbf{x})$. In order to approximate this function a class of functions $\phi(\cdot)$ called Radial Basis

V. Latorre (✉)
University "Sapienza" of Rome, Rome, Italy
e-mail: latorre@dis.uniroma1.it

D. Y. Gao
University of Ballarat, Ballarat, Australia
e-mail: d.gao@ballarat.edu.au

Functions are used. In general the approximating function $g(\cdot)$ has the following form:

$$g(\mathbf{x}) = \sum_{i=1}^{N} w_i \phi(\|\mathbf{x} - \mathbf{c}_i\|), \tag{2}$$

where N is the number of units used to approximate the function, or neurons of the network, $\mathbf{w}$ is the vector with components w_i for $i = 1, \ldots, N$ that is the vector of the weights associated with the connections between the units, and $\mathbf{c}_i \in \mathbb{R}^n$ for $i = 1, \ldots, N$ are the centers of the RBFs. It was shown in [1] that the Eq. (2) presents a universal approximation to general nonlinear functions. An associated optimization problem should be solved in order to train a RBF NN. In this problem, the objective function is formed by two terms: (1) a measure of the approximation error via the least squares; (2) a regularization term used to give to the objective function some desired proprieties such as coercivity. In general the objective function is given in the form of

$$E(\mathbf{w}, \mathbf{c}) = \frac{1}{2} \sum_{p=1}^{P} \sum_{i=1}^{N} (w_i \phi(\mathbf{c}_i) - y_p)^2 + \beta_1 \|\mathbf{w}\|^2, \tag{3}$$

where β_1 is a parameter chosen before the optimization that weights the regularization coefficient. This parameter is always positive and close to zero.

There are mainly two kinds of strategies to train a RBF NN:

1. Supervised selection of weights. In this strategy the center of the RBF functions are set at random or by using a particular strategy, for example by using clustering methods, and then minimize the error function (3) with respect to the weights w_i. In this case the problem is convex;
2. Supervised selection of weights and centers. For this strategy the centers are not set before the optimization, but they are part of the variables that can be chosen for the optimization. In this case the error function becomes:

$$E(\mathbf{w}, \mathbf{c}) = \frac{1}{2} \sum_{p=1}^{P} \sum_{i=1}^{N} (w_i \phi(\mathbf{c}_i) - y^p)^2 + \frac{1}{2} \beta_1 \|\mathbf{w}\|^2 + \frac{1}{2} \beta_2 \sum_{i=1}^{N} \sum_{j=1}^{n} c_{ji}^2. \tag{4}$$

The second strategy seems to create Neural Networks with better approximation capabilities than the ones created by the first strategy, but problem (4) is not convex in $\mathbf{c}$. This characteristic, added to the aspect that the problem has a high number of local minima, determines that the training of a RBF NN with the second strategy can be difficult. One strategy to minimize the nonconvex objective function (4) is the so-called decomposition techniques (see [2]). However, it is difficult to find global optimal solution by traditional direct approaches.

2 Canonical Duality Theory

With the canonical dual theory developed in [3], it is possible to formulate a perfect dual problem in the sense that there is no duality gap and the associated triality theory can be used to identify both global and local optimal solutions. In order to demonstrate the application of this theory, let us consider the following nonconvex minimization problem:

$$(P): \quad \min_{\mathbf{z} \in \mathcal{Z}_a} \left\{ P(\mathbf{z}) = \frac{1}{2} \langle \mathbf{z}, A\mathbf{z} \rangle - \langle \mathbf{z}, \mathbf{f} \rangle + W(\mathbf{z}) \right\}, \tag{5}$$

where $A \in \mathbb{R}^{n \times n}$ is a given symmetric indefinite matrix, $\mathbf{f} \in \mathbb{R}^n$ is a given vector, $\langle \mathbf{v}, \mathbf{v}^* \rangle$ denotes the bilinear form between $\mathbf{v}$ and its dual variable $\mathbf{v}^*$, $W(\mathbf{z})$ is a general non convex function, and $\mathcal{Z}_a \subset \mathbb{R}^n$ is the feasible space for the vector $\mathbf{z}$.

The key step of the canonical dual transformation is to choose a nonlinear operator $\xi = \Lambda(\mathbf{z}) : \mathcal{Z}_a \to \mathcal{E}_a$ and a canonical function $V : \mathcal{E}_a \to \mathbb{R}$ such that the nonconvex function $W(\mathbf{z})$ can be expressed in the canonical form $W(\mathbf{z}) = V(\Lambda(\mathbf{z}))$. A typical example for the nonconvex function $W(\mathbf{z})$ could be the so-called double-well energy

$$W(\mathbf{z}) = \frac{1}{2} \left(\frac{1}{2} \|\mathbf{z}\|^2 - \lambda \right)^2, \tag{6}$$

which has extensive applications in mathematical physics [3] and network optimization (see [4]). In this case the primal problem is

$$P(\mathbf{z}) = \frac{1}{2} \left(\frac{1}{2} \|\mathbf{z}\|^2 - \lambda \right)^2 + \frac{1}{2} \mathbf{z}^T A \mathbf{z} - \mathbf{f}^T \mathbf{z}. \tag{7}$$

For this example, we can simply choose $\xi = \frac{1}{2} \|\mathbf{z}\|^2$ and $V(\xi) = \frac{1}{2}(\xi - \lambda)$. Therefore, the duality mapping $\sigma = V'(\xi)$ from $\mathcal{E}_a = \{\xi \in \mathbb{R} | \xi \geq 0\}$ to its range $\mathcal{E}_a^* = \{\sigma \in \mathbb{R} | \sigma \geq -\lambda\}$ is invertible and the conjugate function $V^*(\sigma)$ can be defined by the Legendre transformation

$$V^*(\sigma) = \text{sta}\{\langle \xi, \sigma \rangle - V(\xi) | \xi \in \mathcal{E}_a\} = \frac{1}{2}\sigma^2 + \lambda\sigma, \tag{8}$$

where the notation $\text{sta}\{g(\xi) | \xi \in \mathcal{E}_a\}$ stands for stationary point of the function $g(\xi)$ on $\mathcal{E}_a$. Therefore, by using the equality $W(\mathbf{z}) = \langle \Lambda(\mathbf{z}), \sigma \rangle - V^*(\sigma)$, the nonconvex function $P(\mathbf{z})$ can be written in the form of the so-called total complementarity function $\Xi(\mathbf{z}, \sigma)$ (see [3])

$$\Xi(\mathbf{z}, \sigma) = \frac{1}{2} \|\mathbf{z}\|^2 \sigma - V^*(\sigma) + \mathbf{z}^T A \mathbf{z} - \mathbf{f}^T \mathbf{z} = \frac{1}{2} \mathbf{z}^T G(\sigma) \mathbf{z} - \sigma - \lambda - \mathbf{f}^T \mathbf{z}, \tag{9}$$

where $G(\sigma) = \sigma I + A$. From this total complementary function, the canonical dual function can be defined by

$$P^d(\sigma) = \text{sta}\{\Xi(\mathbf{z}, \sigma) | \mathbf{z} \in \mathcal{Z}_a\} \tag{10}$$

By the stationary condition $\nabla_{\mathbf{z}}\Xi = G(\sigma)\mathbf{z} - \mathbf{f} = 0$, then on the dual feasible space

$$\mathcal{S}_a = \{\sigma \in \mathcal{E}_a^* | \ \det G(\sigma) \neq 0\},$$

the problem dual to (P) can be formulated as

$$(P^d): \quad \max\left\{P^d(\sigma) = -\frac{1}{2}\mathbf{f}^T G(\sigma)^{-1}\mathbf{f} - V^*(\sigma) \mid \sigma \in \mathcal{S}_a\right\}. \tag{11}$$

Theorem 1 (Complementarity-dual principle [3]) *Problem (P^d) is canonically dual to (P) in the sense that if $(\bar{z}, \bar{\sigma})$ is a critical point of $\Xi(z, \sigma)$, then $\bar{z}$ is a feasible solution of (P), $\bar{\sigma}$ is a feasible solution of (P^d) and*

$$P(\bar{z}) = \Xi(\bar{z}, \bar{\sigma}) = P^d(\bar{\sigma}). \tag{12}$$

Theorem 2 (Analytic solution to (P) [3]). *If $\bar{\sigma} \in \mathcal{S}_a$ is a solution of (P^d), then*

$$\bar{z} = G(\bar{\sigma})^{-1}\mathbf{f} \tag{13}$$

is a feasible solution of (P) and $P(\bar{z}) = P^d(\bar{\sigma})$. Conversely if $\bar{z}$ is a solution of (P), it must be in the form of (13) for certain solutions $\bar{\sigma}$ of (P^d).

In order to study extremalty conditions of the general analytic solution, we define the two following subsets of $\mathcal{S}_a$:

$$\mathcal{S}_a^+ = \{\sigma \in \mathcal{S}_a \mid G(\sigma) \succ 0\}, \ \ \mathcal{S}_a^- = \{\sigma \in \mathcal{S}_a \mid G(\sigma) \prec 0\}.$$

Theorem 3 (Triality theory [3, 5]). *Suppose that $\bar{\sigma}$ is a critical point of P^d and that $\bar{x} = G(\bar{\sigma})^{-1}\mathbf{f}$. If $\bar{\sigma} \in \mathcal{S}_a^+$, then $\bar{z}$ is a global minimizer of (P), $\bar{\sigma}$ is a global maximizer of (P^d), and*

$$\min_{\mathbf{z} \in \mathcal{Z}_a} P(\bar{z}) = \Xi(\bar{z}, \bar{\sigma}) = \max_{\sigma \in \mathcal{S}_a^+} P^d(\bar{\sigma}) \tag{14}$$

If $\bar{\sigma} \in \mathcal{S}_a^-$ and Problem (P) has the same dimension as (P^d), then on a neighborhood $\mathcal{Z}_0 \times \mathcal{S}_0 \subset \mathcal{Z}_a \times \mathcal{S}_a^-$ of $(\bar{z}, \bar{\sigma})$, we have either

$$\min_{\mathbf{z} \in \mathcal{Z}_0} P(\bar{z}) = \Xi(\bar{z}, \bar{\sigma}) = \min_{\sigma \in \mathcal{S}_0} P^d(\bar{\sigma}) \tag{15}$$

or

$$\max_{\mathbf{z} \in \mathcal{Z}_0} P(\bar{z}) = \Xi(\bar{z}, \bar{\sigma}) = \max_{\sigma \in \mathcal{S}_0} P^d(\bar{\sigma}) \tag{16}$$

This triality theory can be used to identify both global and local extrema of the nonconvex problem (P). Extensive applications of the canonical duality theory have been given in fields of computational biology [6], mathematical physics [7], and discrete and network optimization [8, 9].

3 Dual Formulation for Radial Basis Neural Networks

We now turn our attention to the proposed problem (4). In our application we consider only one variable $\mathbf{c}$ since the problem is convex in respect to $\mathbf{w}$. The type of radial basis function we consider is the Gaussian

$$\phi(\mathbf{c}) = \exp\left\{-\frac{\|\mathbf{x} - \mathbf{c}\|^2}{2\alpha^2}\right\}, \tag{17}$$

where α is a parameter belonging to $\mathbb{R}$ greater than zero. The problem becomes

$$P(\mathbf{c}) = \frac{1}{2}\sum_{p=1}^{P}\left(\sum_{i=1}^{N} w_i \exp\left\{-\frac{\|x_p - c_i\|^2}{2\alpha^2}\right\} - y_p\right)^2 + \frac{1}{2}\beta\langle\mathbf{c};\mathbf{c}\rangle - \langle\mathbf{c};\mathbf{f}\rangle, \tag{18}$$

where $\mathbf{f} = \{f_{ji}\} \in \mathbb{R}^{n \times N}$ is a given perturbation and $\langle\mathbf{c};\mathbf{c}\rangle = \sum_{j=1}^{n}\sum_{i=1}^{N} c_{ji}^2$. In order to obtain the dual canonical function we need to use the sequential canonical dual transformation introduced in [3] and the following definitions:

- $\xi_p = \Lambda_p(c) = \sum_{i=1}^{N} w_i \exp\{-\frac{\|x_p - c_i\|^2}{2\alpha^2}\}$ for $p = 1,\ldots P$ as the first level nonlinear operator, with $\mathcal{Z}_a$ being $\mathbb{R}^{\mathbb{P}}$;
- $\sigma_p = \xi_p - y_p$ the first level dual variables, with $\mathcal{S}_a := \{\sigma \in \mathbb{R}^{\mathrm{P}}| -y_{\mathrm{p}} \leq \sigma_{\mathrm{P}} \leq \sum_{i=1}^{\mathrm{N}} w_i - y_{\mathrm{p}}$, if $y_p \geq 0$ or $\sum_{i=1}^{N} w_i - y_p \leq \sigma \leq -y_p$ if $y_p < 0$ for $p = 1,\ldots P\}$;
- $\epsilon_{ip} = \|x_p - c_i\|^2$ as the second level nonlinear operator with $\mathcal{E}_a := \{\epsilon_{ip} \in \mathbb{R}|\epsilon \geq 0 \;\forall i = 1,\ldots,\mathrm{N}, \; p = 1,\ldots,\mathrm{P}\}$;
- $\tau_{ip} = \frac{w_i}{2\alpha^2}\exp\{-\frac{\epsilon}{2\alpha^2}\}$ as the second level dual variables with $\mathcal{T}_{a:} = \{\tau_{ip} \in \mathbb{R}|\tau_{ip} < 0$ if w_i or $\tau_{ip} > 0$ if $w_i < 0$ for $i = 1,\ldots,N, \; p = 1,\ldots,P\}$.

The dual canonical function is

$$P_2^d(\boldsymbol{\tau}, \boldsymbol{\sigma}) = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{n}\left[\frac{\left(f_{ij} + 2\sum_{p=1}^{P} x_{jp}\tau_{ip}\sigma_p\right)^2}{\beta + 2\sum_{p=1}^{P}\tau_{ip}\sigma_p} + \sum_{j=1}^{P} x_{jp}^2\tau_{ip}\sigma_p\right]$$
$$- U^*(\boldsymbol{\tau}, \boldsymbol{\sigma}) - V^*(\boldsymbol{\sigma}), \tag{19}$$

where

$$U^*(\tau, \sigma) = -2\alpha^2 \sum_{i=1}^{N} \sum_{p=1}^{P} \left(\ln\left(\frac{2\alpha^2 \tau_{ip}}{w_i}\right) - 1 \right) \tau_{ip}\sigma_p, \quad V^*(\sigma) = \frac{1}{2}\langle\sigma,\sigma\rangle + \sum_{p=1}^{P} y_p\sigma_p.$$

By the relation $\sigma_p = -2\alpha^2 \sum_{i=1}^{N} \tau_{ip} + y_p \quad p = 1, \ldots, P$, the canonical dual function can be finally written as

$$P^d(\tau) = P_2^d(\tau, \sigma(\tau)). \tag{20}$$

4 Numerical Examples

The numerical examples are in the one dimensional case, with $P = N = n = 1$, $f = 0$, $\alpha = \frac{\sqrt{2}}{2}$ and $\beta = 0.1$. The dual function is:

$$-\frac{1}{2}\frac{(f - 2\tau x(2\alpha^2\tau + y))^2}{\beta - 2\tau(2\alpha^2\tau + y)} - x^2\tau(2\alpha^2\tau + y) - \frac{1}{2}(2\alpha^2\tau + y)^2$$

$$+ (2\alpha^2\tau + y)\left(y - 2\alpha^2\tau\left(\ln\left(\frac{-2\alpha^2\tau}{w}\right) - 1 \right) \right)$$

In Fig. 1 it is possible to see the first example with $x = y = 1$ and $w = 2$. This example has three critical points in the primal and four critical points in the dual, two critical point in S_a^+ and 2 critical points in S_a^-. There is one more critical point in the dual than the primal. This point is called "false dual critical point", it is due to the sequential dual canonical transformation and does not have any corresponding point in the primal problem, unless for a particular configuration of the parameter in input. It always has a fixed value, that is $\tau_f = -\frac{y}{4\alpha^2}$. In the

Fig. 1 Graphs of $P(c)$ (*blue*) and $P^d(\tau)$ (*red*) with $x = y = 1$, $w = 2$
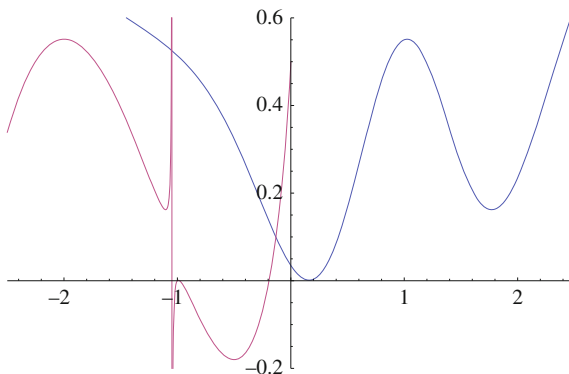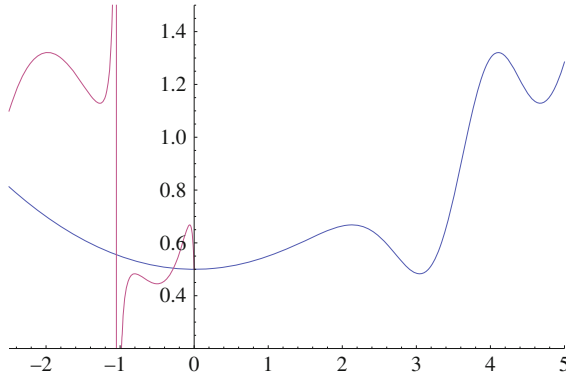
Fig. 2 Graphs of $P(c)$ (*blue*) and $P^d(\tau)$ (*red*) with $x = 4$, $y = 1$, $w = 2$

multidimensional case its position is $\sum_{i=1}^{N} \tau_{ip} = -\frac{y_p}{4\alpha^2}$ for $p = 1, \ldots, P$. The positions and the function values in the critical points are:

- local min $c_1 = 1.7708$, $\tau_1 = -1.1040$, $P(c_1) = P^d(\tau_1) = 0.1622$;
- local max $c_2 = 1.0257$, $\tau_2 = -1.9987$, $P(c_2) = P^d(\tau_2) = 0.5513$;
- global min $c_3 = 0.1616$, $\tau_3 = -0.9903$, $P(c_3) = P^d(\tau_3) = 0.0014$.

All the critical points have the same value of the objective function and the value of the global minimum is in $\mathcal{S}_a^+$. This example respects the general canonical duality framework, apart for the critical point σ_f. In Fig. 2 it is possible to see the second example with $x = 4$, $y = 1$, and $w = 2$. This example has 5 critical point in the primal problem and 6 critical point in the dual, with the sixth critical point in the dual being $\tau_f = -\frac{y}{4\alpha^2} = 0.5$. The critical points are:

- highest local min $c_1 = 4.6721$, $\tau_1 = -1.2730$, $P(c_1) = P^d(\tau_1) = 1.1287$;
- highest local max $c_2 = 4.1062$, $\tau_2 = -1.9776$, $P(c_2) = P^d(\tau_2) = 1.3209$;
- 2nd highest local min $c_3 = 18 \cdot 10^{-6}$, $\tau_3 = -2 \cdot 10^{-7}$, $P(c_3) = P^d(\tau_3) = 0.5$;
- 2nd highest local max $c_4 = 2.1303$, $\tau_4 = -0.6065$, $P(c_4) = P^d(\tau_4) = 0.6681$;
- global min $c_5 = 0.1616$, $\tau_5 = -0.8014$, $P(c_5) = P^d(\tau_5)0.4829$.

This example has four critical points in $\mathcal{S}_a^+$. The second highest local minimum near the frontier of the feasible set for the dual that is $\tau_3 = -2 \cdot 10^{-7}$ is a critical point generated by the regularization term $\frac{1}{2}\beta\|c\|$. This term, even if it important for the properties it gives to the objective function, doesn't express any measure of error of the prediction of the neural network. Also this point will always be near zero because it is the minimum of the function $\frac{1}{2}\beta\|c\|$. For c near to zero the value of $\xi = we^{-\frac{\|x-c\|^2}{2\alpha^2}}$ also is near zero, and same thing happens for $\tau = -\frac{w}{2\alpha^2}e^{-\frac{\|x-c\|^2}{2\alpha^2}}$, in other words this critical point will always be near the boundary and easily found.

Plus in order to have a good prediction it would be preferable if τ is as closer as possible to $-2\alpha^2 y$.

For certain values of the parameter β it can happen that the minimum near $\tau = 0$ is a global minimum of the problem. For example if the value of β in the previous example is changed from 0.1 to 0.12 the point $\tau = -2 \cdot 10^{-7}$ becomes the critical point corresponding to the global minimum in the primal with $P(c) = P^d(\tau) = 0.5$ while the critical point most closer to $\tau = -2\alpha^2 y$ is $\tau = -0.75768$, $P(c) = P^d(\tau) = 0.574697$. In this case the point most closer to $-2\alpha^2 y$ should be chosen as it is the best solution for the original problem. With the two following theorems we are able, for the linear case, to set some conditions on β so that we are sure that there is always a critical point in the interval $G(\tau) > 0 \cap |\tau| > |-\frac{y}{4\alpha^2}|$:

Theorem 4 *Let $\tau_f = -\frac{y}{4\alpha^2}$ the false critical point in the dual problem. then:*

- *if x positive, $x \in \left(0, \sqrt{-2\alpha^2 \ln\left(\frac{y}{2w}\right)}\right)$ or x negative, $x \in \left(-\sqrt{-2\alpha^2 \ln\left(\frac{y}{2w}\right)}, 0\right)$ then τ_f is always a local minimum of $P^d(\sigma)$;*
- *if $x > \sqrt{-2\alpha^2 \ln\left(\frac{y}{2w}\right)}$ or $x < -\sqrt{-2\alpha^2 \ln\left(\frac{y}{2w}\right)}$ then:*

  - *if $\beta > 0$, $x \geq 0$ and $\beta < \frac{y^2 \sqrt{-2\alpha^2 \ln\left(\frac{y}{2w}\right)}}{4\alpha^2 \left(x - \sqrt{-2\alpha^2 \ln\left(\frac{y}{2w}\right)}\right)}$ or $x < 0$ and $\beta < \frac{-\sqrt{-2\alpha^2 \ln\left(\frac{y}{2w}\right)}}{x + \sqrt{-2\alpha^2 \ln\left(\frac{y}{2w}\right)}}$, τ_f is a local minimum for the dual problem;*

  - *if $\beta > 0$, $x \geq 0$ and $\beta > \frac{y^2 \sqrt{-2\alpha^2 \ln\left(\frac{y}{2w}\right)}}{4\alpha^2 \left(x - \sqrt{-2\alpha^2 \ln\left(\frac{y}{2w}\right)}\right)}$ or $x < 0$ and $\beta > \frac{-\sqrt{-2\alpha^2 \ln\left(\frac{y}{2w}\right)}}{x + \sqrt{-2\alpha^2 \ln\left(\frac{y}{2w}\right)}}$, τ_f is a local maximum for the dual problem;*

  - *if $\beta > 0$, $x \geq 0$ and $\frac{y^2 \sqrt{-2\alpha^2 \ln\left(\frac{y}{2w}\right)}}{4\alpha^2 \left(x - \sqrt{-2\alpha^2 \ln\left(\frac{y}{2w}\right)}\right)}$ or $x < 0$ and $\beta = \frac{-\sqrt{-2\alpha^2 \ln\left(\frac{y}{2w}\right)}}{x + \sqrt{-2\alpha^2 \ln\left(\frac{y}{2w}\right)}}$, τ_f is an inflection point in which the first order derivative is zero and that corresponds to a a local minimum of the primal problem.*

Theorem 5 *Let $\tau_f = -\frac{y}{4\alpha^2}$ be the false critical point in the dual problem. Then*

- *if τ_f is a local minimum for the dual function, there will be a local maximum in $(|\tau| > |-\frac{y}{4\alpha^2}|) \cap G(\tau) > 0$ that corresponds to a minimum of the primal problem.*
- *if τ_f is a local maximum then:*

  - *there are no critical points in $(|\tau| > |-\frac{y}{4\alpha^2}|) \cap G(\tau) > 0$;*
  - *there is at least one critical point in $(|\tau| < |-\frac{y}{4\alpha^2}|) \cap G(\tau) > 0$.*

By lowering the value of β under a certain threshold it is possible to always assure that there is a local minimum in $G(\tau) > 0 \cap |\tau| > |-\frac{y}{4\alpha^2}|$.

5 Conclusions

With the formulation (20) and the decomposition algorithms [2], it is possible to find a good minimum for the RBF neural network training problem. The principal framework of this kind of algorithms should be to solve a canonical dual problem when it comes to optimize in respect of the centers c, and then apply a minimum square approach when it comes to minimize in respect to the weights w_i. There are still other open topics such as experimenting on other kinds of RBF or producing a dual that considers all the variable together.

References

1. Haykin S (1999) Neural networks, a comprehensive foundation. Prentice-Hall, Prentice
2. Buzzi C, Grippo L, Sciandrone M (2001) Convergent decomposition techniques for training RBF neural networks. Neural Comput 13:1891–1920
3. Gao DY (2000) Duality principles in nonconvex systems: theory, methods, and applications. Kluwer Academic Publishers, Dordrecht Nonconvex Optimization and Its Applications
4. Gao DY, Ruan N, Pardalos PM (2011) Canonical dual solutions to sum of fourth-order polynomials minimization problems with applications to sensor network localization. In: Pardalos PM, Boginski V, Commander C (eds) Sensors: theory, algorithms and applications. Springer, New York
5. Gao DY, Wu CZ (2012) On the triality theory for a quartic polynomial optimization problem. J Ind Manage Optim 8(1):229–242
6. Zhang J, Gao DY, Yearwood J (2011) A novel canonical dual computational approach for prion AGAAAAGA amyloid fibril molecular modeling. J Theor Biol 284:149–157. doi:10.1080/10556788.2011.641125
7. Santos HAFA, Gao DY (2012) Canonical dual finite element method for solving post-buckling problems of a large deformation elastic beam. Int J Nonlinear Mech 47:240–247. doi:10.1016/j.ijnonlinmec.2011.05.012
8. Gao DY (2009) Canonical duality theory: theory, method, and applications in global optimization. Comput Chem 33:1964–1972
9. Gao DY, Watson LT, Easterling DR, Thacker WI, Billups SC (2011) Solving the canonical dual of box- and integer-constrained nonconvex quadratic programs via a deterministic direct search algorithm. Optim Methods Softw. doi:10.1080/10556788.2011.641125

Support Vector Machines for Real Consumer Circuits

Ciccazzo Angelo, Di Pillo Gianni and Vittorio Latorre

Abstract Circuit analysis is an important phase of the circuit production process. This phase should be performed as fast as possible because of the strict temporal constraints in the industrial sector. On the other hand, there is the need of a certain precision and reliability of the analysis. For this reasons there is more and more interest toward surrogate models that are able to perform a reliable analysis in less time. In this work we analyze how a popular surrogate model, the Support Vector Machines (SVM), performs when it is used to approximate the behavior of industrial circuits, provided by ST-Microelectronics, that will be employed in consumer electronics. The SVM are also compared with the surrogate models created with the commercial software currently used by ST-Microelectronics for this kind of applications.

1 Introduction

Circuit design is a phase of circuit manufacturing process. A circuit must be tested in several ways in order to be ready for massive production. These tests determine if a circuit satisfies its specifications in different possible scenarios. If the circuit

C. Angelo (✉)
ST-Microeletronics Catania, Geneva, Switzerland
e-mail: angelo.ciccazzo@st.com

D. P. Gianni · V. Latorre
Department of Computer, Controls and Management Engineering A. Ruberti,
Sapienza University of Rome, Rome, Italy
e-mail: dipillo@dis.uniroma1.it

V. Latorre
e-mail: latorre@dis.uniroma1.it

does not satisfy its specifications, it can bring to circuit malfunctions or circuit failures. These tests are performed with computer simulations. The computer simulates the behavior of the circuit for the time needed to calculate its performances.

The time needed to perform a simulation changes according to the complexity of the circuit, it can take from a minute to several days to complete it. Generally speaking, these simulations can be considered costly. In industrial design there are strict temporal limitation due to the time to market. So the faster a circuit analysis is performed, the better it is. For this reason there is a great interest in the use of surrogate models to tackle the problem of circuit design. Surrogate models are mathematical models that use couples of input-output samples to create a regression function that approximates the relation between the input parameters and the output performances.

In literature, several works about surrogate models applied to circuit simulations are proposed. Especially Kriging models [1, 2] and response surface methodologies [3].

In this paper we use the Support Vector Machine (SVM) [4] to create surrogate models based on computer simulations and then compare them with one of the most used mathematical models for this kind of applications, the Response Surface Methodologies (RSM) [5], realized with a commercial software widely used in the industrial sector, Wicked [6]. The simulations are based on circuits designed and produced in the industrial sector by ST-Microelectronics. These circuits will become parts of consumer technology devices.

The principal aim of this work is to understand if the SVM are useful in modeling actual industrial circuits in order to shorten the time needed to perform a circuit analysis. To this aim the SVM are tested on real circuits and compared with a software suite expressly created for circuit analysis. We are also interested in using as few simulations as possible to obtain a good model, so we perform other tests reducing the number of samples in the training in order to understand how many simulations are needed to obtain a reasonable precision.

2 Surrogate Models

Surrogate models are used in a plethora of different application and problems [4]. The regression problem consists in: Given a set of P samples $T = \{(x^p, y^p), x^p \in \Re^n, y^p \in \Re, \quad p = 1, ..., P\}$ we are interested in finding the regression function $f : \Re^n \to \Re$ that approximates the relation between the vector of input variables $x \in \Re^n$ and the output $y \in \Re$.

Support Vector machines are one of the most used models for regression. The SVM calculates the regression function according to the following formula:

$$y(x) = \sum_{p=1}^{P}((\hat{\lambda}^p)^* - (\lambda^p)^*)k(x, x^p) + b^*, \qquad (1)$$

where the $(\hat{\lambda}^p)^*$, $(\lambda^p)^*$ and b^* are obtained by solving the following optimization problem:

$$\min \Gamma(\lambda, \hat{\lambda}) = \frac{1}{2}\sum_{p=1}^{P}\sum_{q=1}^{P}(\hat{\lambda}^p - \lambda^p)(\hat{\lambda}^q - \lambda^q)k(x^p, x^q)$$

$$- \sum_{p=1}^{P}(\hat{\lambda}^p - \lambda^p)y^p + \epsilon\sum_{p=1}^{P}(\hat{\lambda}^p + \lambda^p)$$

$$(2)$$

$$\sum_{p=1}^{P}(\hat{\lambda}^p - \lambda^p) = 0,$$

$$0 \le \lambda^p \le C \quad p = 1, \ldots, P,$$

$$0 \le \hat{\lambda}^p \le C \quad p = 1, \ldots, P.$$

In the objective function there is the kernel function $k(\cdot, \cdot)$. We use the Gaussian kernel described by the following formula:

$$k(x^p, x^q) = \exp(-\sigma\|x^p - x^q\|^2). \qquad (3)$$

We note in problem (2) the presence of a parameter C. This parameter controls the trade-off between over-fitting and knowledge in the model. We also note that in Eq. (3) the presence of a parameter σ that influences the kernel function. The performances of the SVM greatly depends on the choice of these parameters. In order to obtain the best prediction as possible, we adopt a grid search cross-validation strategy. We use the Mean square error (MSE) to measure the quality of the prediction:

$$MSE := \frac{1}{P}\sum_{i=1}^{P}(f_i - y_i)^2,$$

where P is the number of samples in the set, f_i is the output calculated by the surrogate model and y_i is the real output.

In order to make the prediction less dependent on the training set given in input to the SVM, we also adopt a k-fold strategy for cross-validation. In our application we set k = 10, a choice often used in literature. In order to train the support vector machine we use the software LIBSVM [7].

Response Surface Methodologies (RSM) [5] calculate the regression function according to the following formula:

$$y(x) = \beta_0 + \sum_{j=1}^{n} \beta_j x_j, \tag{4}$$

where the $\beta_j, j = 0, \ldots, n$ are the regression parameters calculated by solving the following system of linear equations:

$$y_p = \beta_0 + \sum_{j=1}^{n} \beta_j x_{pj} + \epsilon_p, \quad p = 1, \ldots, P \tag{5}$$

where ϵ_p denotes the measurement error for every sample. This error usually has $E(\epsilon_p) = 0 \quad \forall p = 1, \ldots, P$ and the single ϵ_p are uncorrelated to each other.

If $P > n$, a system with P equations and $n + 1$ unknowns is generated. In this case the regression parameters $\beta_j, j = 0, \ldots, n$ that minimize the sum of square of the errors ϵ_p are chosen, that is:

$$\sum_{p=1}^{P} \epsilon_p^2 = \sum_{p=1}^{P} (y_p - \beta_0 - \sum_{p=1}^{P} \beta_j x_{pj})^2. \tag{6}$$

It is possible to further expand this model by performing the following substitution in Eqs. (4), (5) and (6):

$$x_j = g(x_j)$$

where $g(\cdot)$ is a basis function.

The RSM are realized by the software platform Wicked. This software platform implements four types of basis functions: Polynomial, Radial Basis, Affine plus Radial Basis and Polynomial plus Radial Basis (for further details see [8] Chap. 6).

Wicked chooses automatically the model for the RSM with the function that fits data the best by using the information given in the training set. After performing the model selection, Wicked solves problem (5), generates the surrogate model and tests it on the test set, that is the same of the SVM.

3 Experimental Set Up

In this section we explain the steps taken to obtain the reported results. In circuit design there are three types of parameters:

- Design parameters: determine the sizing of the circuit components;
- Process parameters: these parameters represents the statistical variations that could be for example due to the fluctuations in the manufacturing process of the circuit;
- Operating parameters: these parameters describe the environment where the circuit might work.

A high number of parameters in input causes the so called "Curse of dimen-sionality", so the best choice is to use the less parameters as possible without losing information on the phenomenon we want to model. In order to handle this problem, the sensitivity analysis method is used in our experimentation.

The performances are the output of the simulation that the SVM and RSM will approximate they can represent the delay between two waveforms, the duration of an impulse, the variation of the voltage due to this impulse, the slope of the change of voltage etc.

After defining the number of parameters, the sets in input to the surrogate models are generated. Following many works presented in literature [9] we choose the Latin Hypercube Sampling as Design Of Experiment scheme for generating the points in the training and test set for both the SVM and Wicked RSM. The performances of these working points are calculated using the ELDO simulator. Then the data is normalized before the training phase.

In order to determine the number of samples needed to obtain the desired precision, after the first test, we decrease the number of samples in order to see how the two models would fare. This is an important test, because the less sim-ulations are used to realize a surrogate models, the best it is from a design point of view. Both the SVM and RSM are trained on the same training sets and every generated model is tested on the same test set, comprised of 500 samples not utilized in the in the training and model selection phase.

The reported results on the test set will not only take into account the MSE on the denormalized performances, but also the coefficient of determination:
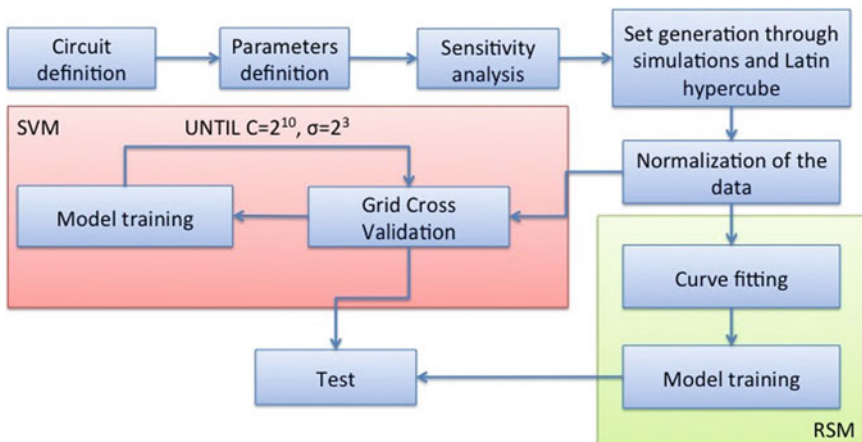
Fig. 1 Flow chart of the experimental set up

$$R^2 := \frac{\sigma_{f,y}^2}{\sigma_f * \sigma_y},$$

where $\sigma_{f,y}$ is the covariance between the calculated output f and the real output y, σ_f and σ_y are the variance of the calculate output and the real output respectively (Fig. 1).

4 Results

In this section we report the results of the experiments and the effect of lowering the number of samples in input to the surrogate models.

The digital to analog converter (DAC) circuit is used for voltage generation and timing in input of cells of flash memory products. The architecture of the DAC circuit is shown in Fig. 2. This circuit has originally in input 54 parameters. After a screening of the parameters performed with a sensitivity analysis, the number of parameters is reduced to 28. In Table 1 we report the results with 28 parameters in input with 500 samples for test for both the SVM and the RSM.

By looking at the experiment with 1,000 samples in training set, we note that the overall behavior of the two methods yields similar results with the SVM being slightly superior. In particular the SVM have an edge when it comes to predict the performances GAIN, SLOPE while the RSM has better results for the VXP-6. We
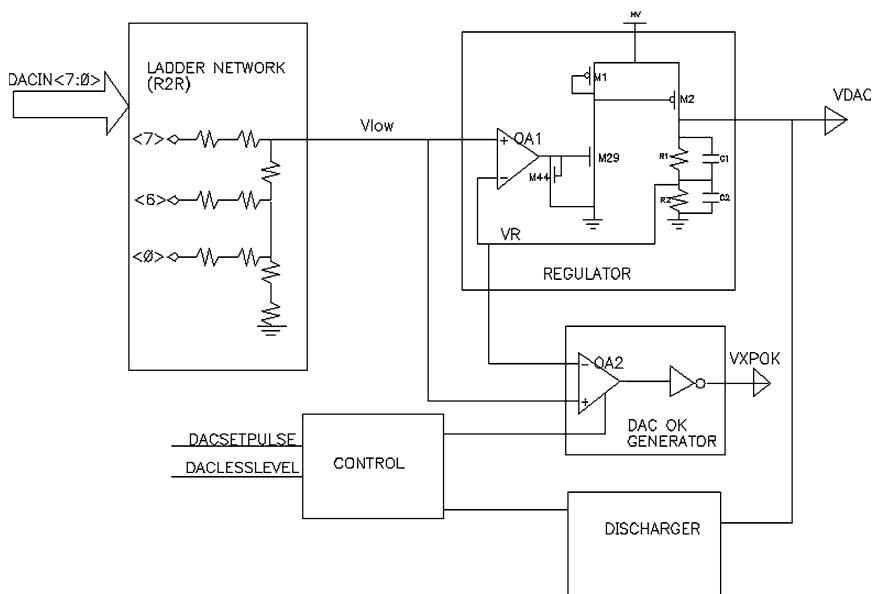
Fig. 2 DAC architecture

Table 1 MSE and R^2 results for the 6 performances diminishing the number of samples in the training on the same test set of 500 samples

ERR	Gain		Slope		VXP-3O3	
	SVM	RSM	SVM	RSM	SVM	RSM
mse	5.09E–06	8.25E–06	1.75E–05	2.47E–05	2.68E–04	2.42E–04
R^2	0.86	0.773	0.937	0.909	9.47E–01	9.52E–01
mse	7.66E–06	1.18E–05	2.44E–05	3.46E–05	3.11E–04	3.06E–04
R^2	0.787	0.673	0.911	0.874	0.94	0.94
mse	1.28E–05	1.31E–05	3.81E–05	3.86E–05	2.97E–04	3.80E–04
R^2	0.659	0.636	0.863	0.859	0.941	0.933
mse	1.36E–05	1.89E–05	4.25E–05	5.46E–05	3.40E–04	4.70E–04
R^2	0.621	0.489	0.845	0.805	0.933	0.913
mse	1.74E–05	2.07E–05	4.80E–05	6.14E–05	3.90E–04	4.70E–04
R^2	0.517	0.456	0.824	0.788	0.925	0.906
mse	2.11E–05	2.26E–05	6.66E–05	8.57E–05	6.22E–04	8.48E–04
R^2	0.414	0.387	0.758	0.704	0.88	0.832
mse	3.69E–05	2.70E–05	8.21E–05	9.69E–05	7.86E–04	8.96E–04
R^2	0.21	0.252	0.701	0.671	0.847	0.822
ERR	VXP-6		VXPOK2		VXP-9	
	SVM	RSM	SVM	RSM	SVM	RSM
mse	4.88E–05	3.41E–05	5.50E–03	6.04E–03	1.56E–04	1.72E–04
R^2	0.987	0.991	0.933	0.926	0.963	0.96
mse	5.43E–05	5.81E–05	8.34E–03	1.00E–02	2.13E–04	1.72E–04
R^2	0.986	0.985	0.897	0.878	0.95	0.96
mse	7.27E–05	7.04E–05	1.27E–02	1.29E–02	3.11E–04	4.12E–04
R^2	0.981	0.981	0.854	0.842	0.926	0.903
mse	1.01E–04	8.62E–05	1.29E–02	1.39E–02	4.19E–04	4.21E–04
R^2	0.974	0.978	0.843	0.83	0.903	0.903
mse	1.32E–04	1.03E–04	1.82E–02	1.93E–02	5.62E–04	5.55E–04
R^2	0.966	0.973	0.792	0.764	0.869	0.87
mse	2.04E–04	1.97E–04	2.18E–02	2.69E–02	8.72E–04	9.03E–04
R^2	0.951	0.949	0.746	0.69	0.797	0.791
mse	3.49E–04	3.11E–04	4.16E–02	4.48E–02	1.08E–03	8.47E–04
R^2	0.915	0.931	0.52	0.476	0.747	0.806

also note that the majority of the models have satisfying results with the exception of the surrogate models for the GAIN performance that nevertheless yield good results.

As concerns the experiments with a decreasing number of samples in the training set, we notice that the two methods have comparable results with the SVM being a little better, especially when it comes to predicting the following performances: GAIN, SLOPE, VXP-303 and VXPOK2, while the RSM is still better in the prediction of the VXP-6.

The two models still have a good prediction rate even after decreasing the number of samples. We note a steady decrease of the surrogate models precision.

In particular if we take as example the results obtained with only 35 samples in input, we note that the surrogate models for the performance VXP-6 still yield satisfying results, while there are still good results for other four performances and bad results for the remaining one.

From these results we can state that the SVM seems to yield sightly better results than the RSM and that the surrogate models are able to give, in the majority of cases, more than reasonable results, even with a limited number of samples in the training set.

5 Conclusions

The aim of this work is to apply surrogate models to real industrial circuits, in order to find a method to analyze these circuits without resorting to a heavy use of costly circuit simulations. We confronted the SVM with the benchmark normally used in the industrial applications, Response Surface Methodologies and obtained comparable results. An important result is that the surrogate models are able to get reasonable prediction for such a complex problem as approximation of circuit simulations even with a limited number of samples in the training set.

The results of our experimentation seems to indicate that SVM is a valid surrogate model for real industrial circuits approximation and it can be considered a valuable tool for applications regarding the circuit analysis.

References

1. Gorissen D, De Tommasi L, Crombecq K, Dhaene T (2009) Sequential modeling of a low noise amplifier with neural networks and active learning. Neural Comput Appl 18: 485–494 (Springer, London)
2. Gorrisen D, De Tommasi L, Hendrickx W, Croon J, Dhaene T (2008) RF circuit block modeling via Kriging surrogates. Microwaves, radar and wireless communications. MIKON 2008. In: 17th international conference. pp 1–4
3. Hailong Y, Maofeng Y, Dan W, Zinzhang J (2009) Kriging model combined with latin hypercube sampling for surrogate modeling of analog integrated circuit performance. Quality of Electronic Design, ISQED. IEEE Press, New York.
4. Schölkopf B, Smola AJ (2002) Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT Press, Cambridge
5. Myers RH, Montgomery DC, Anderson-Cook CM (2009) Response surface methodology: process and product optimization using designed experiments. Wiley, London
6. WiCkeD, a Tool Suite for nominal and statistical custom IC design. MunEDA inc, http://www.muneda.com/Home
7. Chang CC, Lin CJ (2011) LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol 2(3): 27:1–27:27 (ACM, New York)
8. WiCkeD (2011) Manual 6.4-3. MunEDA inc
9. Zhang P, Breitkopf P, Knopf-Lenoir C, Zhang W (2011) Diffuse response surface model based on moving Latin hypercube patterns for reliability-based design optimization of

ultrahigh strength steel NC milling parameters. structural and multidisciplinary optimization. Springer, Berlin/Heidelberg, pp 613–628

10. Bo L, Dixian Z, Reynaert P, Gielen GGE (2011) Synthesis of integrated passive components for high-frequency RF ICs based on evolutionary computation and machine learning techniques. Computer-aided design of integrated circuits and systems, IEEE Transactions on. IEEE Press, New York

11. Smith TF, Waterman MS (1981) Identification of common molecular subsequences. J Mol Biol 147:195–197

12. Hsu CW, Chang CC, Lin CJ (2008) A practical guide to support vector classification. Department of Computer Science and Information Engineering, National Taiwan University

A Triangulation Method for Unorganized Points Cloud Based on Ball Expanding

Qiang Zhang, Nan Wang, Dongsheng Zhou and Xiaopeng Wei

Abstract As an important research subject of the CAD, computer geometry, reverse engineering and other areas. The triangulation of the unorganized data has great significance in theory and the practical. In this paper, we introduce the current mainstream methods of triangulation that based on 3D points cloud data as well as the hash tables and put forward a direct triangulations method which based on a ball expanding.

Keywords Points cloud · Direct triangulations · Hash tables · Ball expanding

1 Introduction

With the advent of the information society, the 3D geometric model has become an important data representation whether in scientific or engineering applications. The triangulation for the scattered data points in R^3 has received more and more attention with the development of science and technology. According to the difference of the points, the triangulation problem can be divided into 2D triangulation and 3D triangulation [1]. The research of the triangulation in plane domain is already quite mature, so we will not go into details in this paper. At present, the 3D triangulation can be divided into two kinds: projection method and the spatial direct triangulation. The projection method is to project the scatted points to the 2D plane and do the triangulations in 2D, then split the triangular mesh after map back to the space. The essence of this method return to the plane domain triangulation problem, and the operation of projected will lead to the distortion of the triangular mesh as well as increase the complexity of the triangulations. For the

Q. Zhang (✉) · N. Wang · D. Zhou · X. Wei
Key Laboratory of Advanced Design and Intelligent Computing, Dalian University,
Ministry of Education, Dalian, China
e-mail: zhangq@dlu.edu.cn

complex model, this method can't obtain a satisfactory result. Therefore, the research of the spatial direct triangulation is needed.

2 Related Works

The more recent exhaustive overviews of 3D triangulation methods are presented in [1, 2]. The most common triangulation algorithms can be divided into two main groups: Mesh growing method and Voronoi-based method.

2.1 Mesh Growing Method

Mesh growing method generate the surface from a seed triangle and grow the meshes by using some criteria. Recent years, a number of algorithms have been proposed. In [2], Bernardini et al. put forward the Ball Pivoting algorithm. This method can afford a correct triangulation for the point cloud as well as the noise data, but it requires normal of the point and is not easy to get the appropriate radius. Huang and Menq [3] proposed an algorithm which firstly search the K-nearest neighbors of the two points of the front edge, then delete the points which are illegal. In [4] Lin et al. finds some shortcomings of this method. In 2004, Lin et al. introduced the Intrinsic Property Driven (IPD) algorithm [4], which overcomes the shortcomings in [3] by improves the way of looking for the point. Li et al. improved the Priority Driven in [5] by evaluating shape changes from the estimation of the original surface in 2009. He proved that the triangulation speed of the method was higher than Ball Pivoting and the Cocone. In [6] Angelo et al. put forward a mesh growing algorithm using the Gabriel 2-Simplex criterion. From his experiments result, the quality of the surface is better than Ball Pivoting and the Cocone family in the tessellation of noise points cloud. In order to improve the robust to the area which is not sufficiently sampled, he improved the algorithm in [7].

2.2 Voronoi-Based Method

Voronoi-based method is according to Delaunay tetrahedralization. Edelsbrunner [8] proposed the concept of α-shape, its principle is to delete the edge, triangular and tetrahedral which the radius of the circumscribed circle or circumsphere is greater than and so as to get the reconstructed surface. Recent years, the most famous method is the Crust because of its theoretical guarantees proposed by Amenta et al. in [9]. Later they promoted it to surface reconstruction in 3D space in [10] and [11]. In 2002, Amenta et al. proposed the Cocone in [12] to reduce the time and memory consumption of Crust. Chang et al. [13] pointed that the Cocone

and Crust have a same shortcoming, that the theoretical guarantee of obtaining a correct reconstruction is only possible as long as the point cloud is well sampled. In 2003 and 2006, Dey and Goswami proposed the Tight Cocone [14] and the Robust Cocone [15], the first one can provide a watertight closed surface reconstruction and the second can deal with noisy data. The Super Cocone in [16] which proposed by Dey et al. is good at processing of large scale points cloud. A further evolution of the Crust is the Power Crust proposed by Amenta et al. in [17] which can generate a watertight mesh for any point cloud, but sharp edges or noisy data can result in defectiveness. In [18], Ou Yang et al. introduced a method a method of reconstruction of triangular surfaces via clustering of Delaunay spheres.

3 The Triangulation Algorithm

The triangle mesh construction method in this paper is based on growth method. The specific process as shown in Figs. 1, 2:

3.1 Spatial Index

In this paper we use Hash table (shown in Fig. 2) to build the spatial index of the point. It divides the data set into many cuboids. We store the three-dimensional

Fig. 1 The process of the algorithm
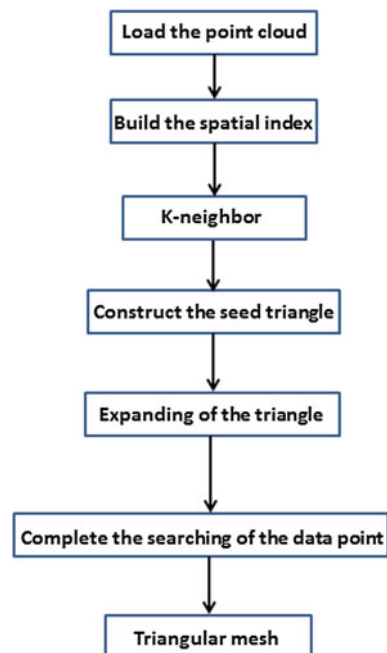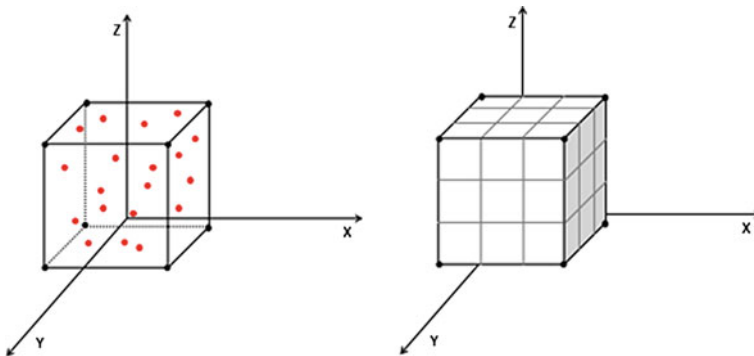
Fig. 2 The hash tables

coordinates of every point in a one-dimensional array. Calculate the maximum value and minimum value of the points cloud along each coordinate X, Y and Z, as $X_{max}, Y_{max}, Z_{max}, X_{min}, Y_{min}, Z_{min}$ Determine the size of the cuboid bounding box, which contain all the points. We can define the size of the cuboid as L_x, L_y, L_z

$$\begin{cases} L_x = X_{max} - X_{min} \\ L_y = Y_{max} - Y_{min} \\ L_z = Z_{max} - Z_{min} \end{cases} \tag{1}$$

The work after calculate the cuboid is the subdivision. Divide the cuboid into many small bounding boxes. The size of the small bounding box is *Medge*

$$Medge = \sqrt[3]{\alpha \frac{L_x L_y L_z}{N}} \tag{2}$$

N is the number of the points cloud, α value according to density of the points cloud, ranges from 10 to 20. The lower the density is, the smaller the α is.

It is clear that the small bounding box of the point is a small cube, so we can get the number of cubes m, n, l along each coordinate X, Y, Z.

$$\begin{cases} m = \text{int}(L_x/Medge) + 1 \\ n = \text{int}(L_y/Medge) + 1 \\ l = \text{int}(L_z/Medge) + 1 \end{cases} \tag{3}$$

And the number of all the cubes is $m \times n \times l$.

Suppose the coordinate value of the current point p is (p_{xi}, p_{yi}, p_{zi}) then

$$\begin{cases} i = (\text{int})(p_{xi} - X_{min})/Medge \\ j = (\text{int})(p_{yi} - Y_{min})/Medge \\ k = (\text{int})(p_{zi} - Z_{min})/Medge \end{cases} \tag{4}$$

Here i, j, k are the point's index values. So, giver one point of the points cloud, the algorithm can easily decide which cubic cell it belongs to.

In KNN search algorithm, we restrict the search range in the 27(include itself) cells in and around the cell the point belongs to. If the index of one point is (i, j, k), the index of the cell around it will be $(i - a, j - b, k - c), a, b, c \in \{-1, 0, 1\}$

Using this method we can get each point's KNNs quickly.

3.2 Construct the Seed Triangle

As is known, a triangle is composed of vertex and edge, so we should construct a vertex firstly, then construct the first edge by this vertex and look for the third vertex in the end. There have a lot of methods about how to construct the seed triangle. In order to improve the efficiency of the algorithm and avoid spending too much time on the judgment, we select the point which the X coordinate value is $X_{\max}$ as the first vertex. When we not get only one point, we can select the one which the Y coordinate value is the largest, and so on. The following are the steps to select the seed triangle:

1. Select the extreme coordinate value point P1 as the first vertex of the seed triangle;
2. Searching for the point P2 within P1′s K-neighbor which the distance to P1 is the nearest;
3. Searching for the third point P3 within the K-neighbor of P1 and P2, make sure that P2P3 + P1P3 as short as possible. The construction will not stop until the P3 is founded, else go to step (2);

3.3 Expand the Triangle

In this paper, the triangle expanding strategy is a ball expanding strategy (Fig. 3 shows an example). We define a minimum surrounded ball which contains the expended edge, searching the expend point by the moving of the center and the

Fig. 3 The mesh and its surrounding ball
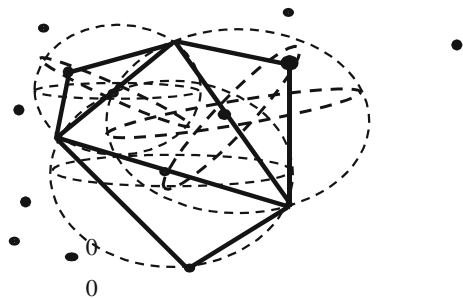
Fig. 4 The expanding edge and its minimum ball
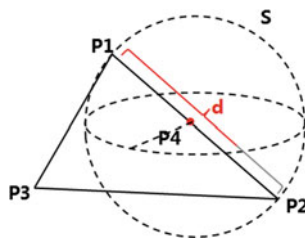
increase of the sphere radius, when the ball touches the third point, the point will be the expanding point. By in this way, we can make the new triangle to be a Delaunay triangle.

3.3.1 The Minimum Ball

When we define the minimum ball in this paper, the ball should contain the two vertices of the expanding edge and is the smallest. Therefore, we can define the center of the sphere as the midpoint of the edge and the radius of the sphere as the half of the length d of the Edge, 1\2d. As shown in Fig. 4: △ P1P2P3 is a existing triangle, P1P2 is the expanding edge, P4 is the midpoint of the edge P1P2, sphere S is the minimum ball that contains the P1P2, the center of the sphere is located in P4 and the radius is 1\2d.

3.3.2 The Moving of the Center

In this algorithm, the expansion of the triangle is based on ball expansion strategy. During the process of expanding, the center of the sphere to be moving along the perpendicular bisector of the expanding edge, and the radius will linear increasing in the same time. In this process, it is ensured that the two vertices of the edge are always on the ball. The expansion point will be founded when the ball touching the third point.

Firstly, we determine the plane which the original belongs to, as shown in Fig. 5, △ P1P2P3. The coordinates of the vertices $P1(x1, y1, z1)$, $P2(x2, y2, z2)$, $P3(x3, y3, z3)$ are known, the equation of the plane that this three vertices established can be defined as $AX + BY + CZ = 1$. There are equations:

Fig. 5 The plane
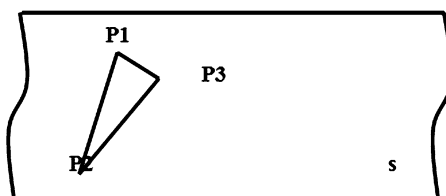
$$\begin{cases} x1 \cdot A + y1 \cdot B + z1 \cdot C = 1 \\ x2 \cdot A + y2 \cdot B + z2 \cdot C = 1 \\ x3 \cdot A + y3 \cdot B + z3 \cdot C = 1 \end{cases} \tag{5}$$

Assumed that $p = \begin{pmatrix} x1, y1, z1 \\ x2, y2, z2 \\ x3, y3, z3 \end{pmatrix}$, so $\begin{pmatrix} A \\ B \\ C \end{pmatrix} = P^{-1}(A, B, C)$ is the normal vector of the plane.

Since the movement of the center of the sphere is moving along in the perpendicular bisector of the edge. We use the parametric equation of the line to define the perpendicular bisector L: $\begin{cases} x = x0 + m \cdot t \\ y = y0 + n \cdot t \\ z = z0 + p \cdot t \end{cases}$. The coordinates of the midpoint O of the expanding edge is $(x0, y0, z0)$, the direction vector of the line L is (m, n, p).

From the above we can see that just need to determine the direction vector can get the equation of the line L. As the point O in line L, take one point D' on L and its coordinate is (x_d, y_d, z_d), D' is always on plane S. Vector $\overrightarrow{OD'}$ perpendicular to the edge $P1P2$, as $\overrightarrow{OD'} \times \overrightarrow{P1P2} = 0$. So we can get equations

$$\begin{cases} x_d \cdot A + y_d \cdot B + Z_d \cdot C = 1 \\ (x_d - x0)(x1 - x2) + (y_d - y0)(y1 - y2) + (z_d - z0)(z1 - z2) = 0 \end{cases}$$

If we define $E = (x1 - x2, y1 - y2, z1 - z2), S = (A, B, C), Q = (x0, y0, z0)$ then the equations can be transformed into:

$$\begin{cases} (x_d, y_d, z_d) \cdot S^T = 1 \\ (x_d, y_d, z_d) \cdot E^T = Q \cdot E^T \end{cases} \tag{6}$$

The solution of the equation is $(x_d, y_d, z_d) = \begin{pmatrix} 1 \\ Q \cdot E^T \end{pmatrix} \cdot \begin{pmatrix} S^T \\ E^T \end{pmatrix}^{-1}$.

After get the coordinates of D', we can obtain the direction of the direction vector $l = (x_d, y_d, z_d) - (x0, y0, z0)$, the unitized result of l is (m, n, p)

Now we can get the movement equation of the center:

$$\begin{cases} x = x0 + m \cdot t \\ y = y0 + n \cdot t \\ z = z0 + p \cdot t \end{cases} \quad t \in (-\infty, +\infty) \tag{7}$$

Here, we also need to determine the positive direction of the movement of the center. We must make sure that the center is moving toward the other side of the triangle, so as to make the new triangle is not exist. In this algorithm, we can take t
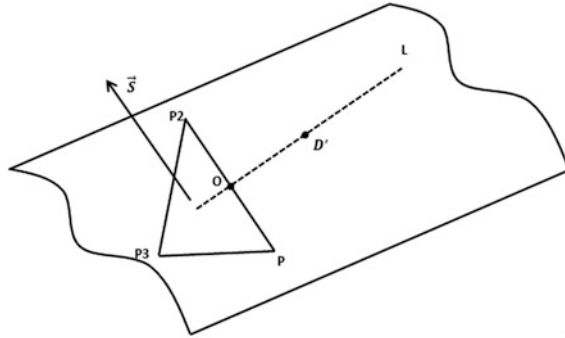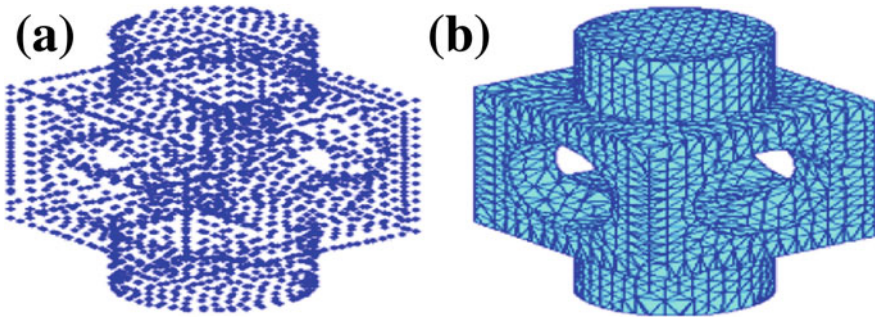
Fig. 6 The movement of the ball

Fig. 7 a Points cloud of the Block, has 2,132 points; **b**, **d** and **e** the growing process of the triangles; **c** the triangulation results, has 4,272 triangles

a positive value and a negative value respectively; it always means take one point on both side of the expanding edge respectively. As shown in the Fig. 7 D'. Now we make cross product $\overrightarrow{P1P2} \times \overrightarrow{P2P3}$ and $\overrightarrow{P1P2} \times \overrightarrow{OD'}$, here get the normal vector of the plane, the direction meet the right hand theorem, as shown in Fig. 6. The two vectors are parallel. So we can make a Vector multiplication to determine the positive direction by judging the positive and negative of the result.

4 Experimental Results

The complexity of the triangulation of this paper is $O(KN)$ because we only search in the K-neighbor each time, K is the number of the points of the K-neighbor, N is the number of the points. In order to prove our algorithm, we realized it in MATLAB. We have used various points cloud. The results are shown in Figs. 7, 8, 9, 10, 11, 12 and Table 1.
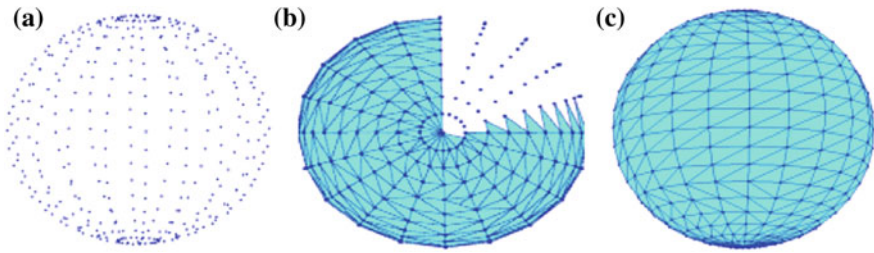
Fig. 8 **a** Points cloud of the ball, has 441 points. The triangulation result has 760 triangles
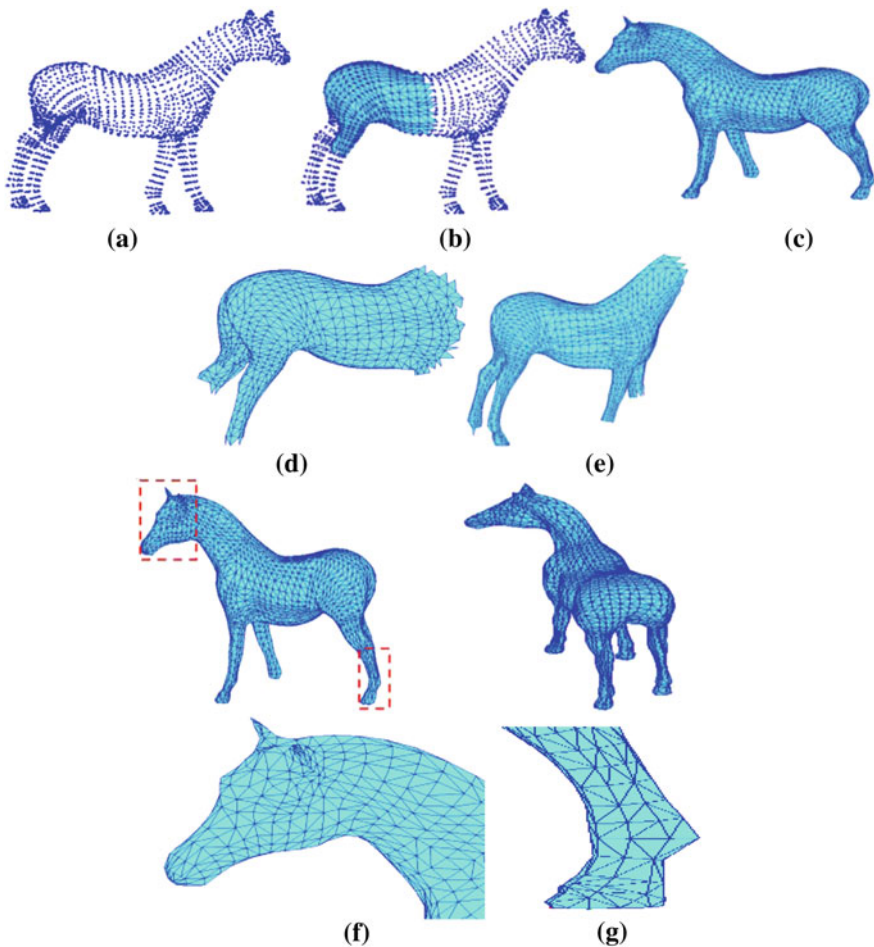
Fig. 9 **a** Points cloud of the Horse, has 2,302 points; **b**, **d** and **e** the growing process of the triangles; **c** the triangulation results, has 4,307 triangles; **f** and **g** the head and leg of the horse. In these areas, the density of the points cloud is large and the transition is drastic. In our results, we can see that there have no thin and narrow triangles at all, and the points in the ears are triangulated perfectly. In our algorithm, we defined that one edge can only be shared by two triangles, so there are no intersection of triangles in the mesh
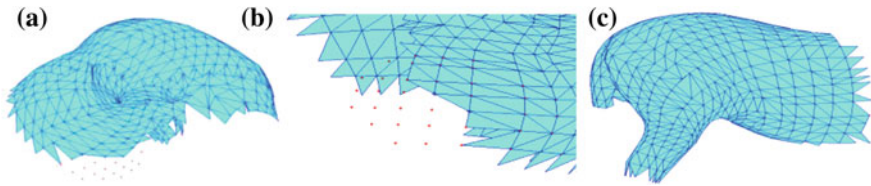
Fig. 10 Shows the growing process. The *red line* is the bounding edge; the red point is the K-nearest neighbors
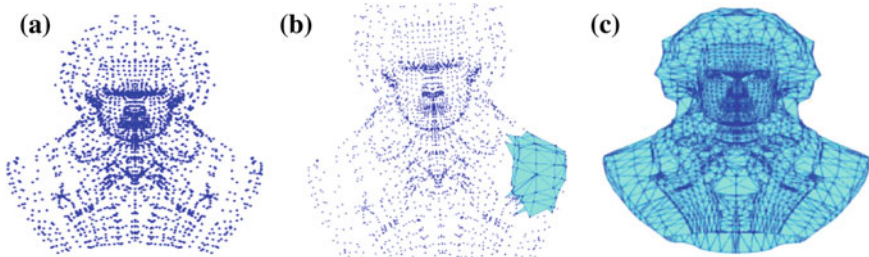
Fig. 11 a The Beethoven points cloud that has 2,655 points; **c** the triangulation results, 4,971 triangles
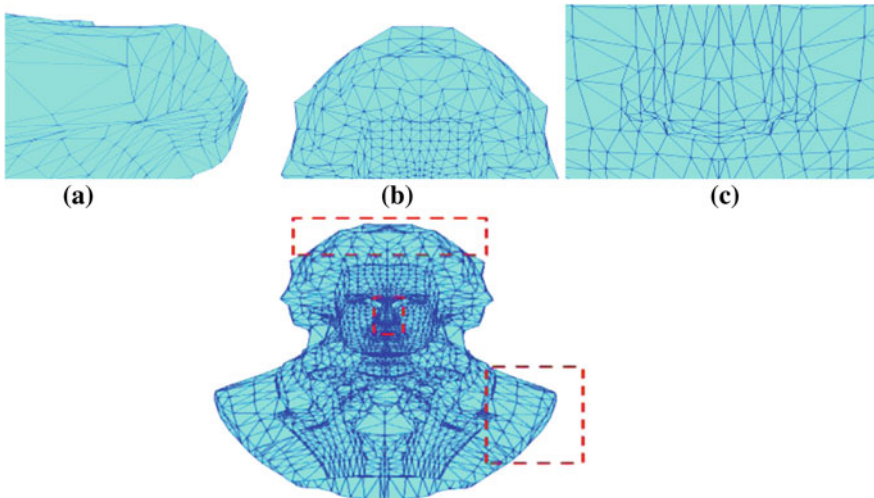
Fig. 12 a The shoulder of Beethoven, in this area the angle of the adjacent triangle is larger, from the result, the meshes formed here are good and no holes exists; **b** the head part, the sharp features are retained and have no intersection of triangles; **c** the nose part of the points cloud, the triangulation of this area should be very complex, because the density of the points is large, the transition is always drastic and there always has some sharp features. As it shown, the result is good, no thin triangles, no holes, no single triangle

Table 1 Experimental results

Model	Points	Triangles	Time (s)
Block	2,132	4,272	6.21
Horse	2,702	4,307	7.02
Beethoven	2,655	4,971	6.88
Ball	411	760	2.15

5 Conclusion

In this paper, we have proposed a new direct triangulation algorithm in 3D, which is computationally and memory efficient. It is always easy to be realized and can obtain a good result. From the experiments, there are no illegal triangles in the mesh. In the future work, we will focus on how to remove the noise data during the triangulation process. When the amount of data is too large, the running time of the algorithm is not satisfactory, and also to be further improved.

References

1. Wei Y, Su X (2008) Novel and fast mapping triangulation algorithm for unorganized pointscloud. Opt Eng 47(11): 117205–117211
2. Bernardini F, Mittleman J, Rushmeier H, Silva C, Taubin G (1999) The ball pivoting algorithm for surface reconstruction. IEEE T Vis Comput Gr 5(4):349–359
3. Huang J, Menq CH (2002) Combinatorial manifold mesh reconstruction and optimization from unorganized points with arbitrary topology. Comput Aided Design 34(2):149–165
4. Lin HW, Tai CL, Wang GJ (2004) A mesh reconstruction algorithm driven by an intrinsic property of point cloud. Comput Aided Design 36(1):1–9
5. Li X, Han CY, Wee WG (2009) On surface reconstruction: a priority driven approach. Comput Aided Design 41(9):626–640
6. Di Angelo L, Di Stefano P, Giaccari L (2011) A new mesh-growing algorithm for surface reconstruction. Comput Aided Design 43(6):639–650
7. Di Angelo L, Giaccari L (2011) A fast algorithm for manifold reconstruction of surfaces. In: Proceedings of the international conference on innovative methods in product design, Venice Italy
8. Edelsbrunner H, Mucke EP (1992) Three-dimensional alpha shapes. In: workshop on volume visualization, Boston, pp 75–82
9. Amenta N, Bern M, Eppstein D (1998) The crust and the beta-skeleton: combinatorial curve reconstruction. GMIP 60(2):125–135

10. Amenta N, Bern M, Kamvysselis M (1998) A new Voronoi-based surface reconstruction algorithm. In: The 25th annual conference on computer graphics and interactive techniques, New York, pp 415–421
11. Amenta N, Bern M (1999) Surface reconstruction by Voronoi filtering. DCG 22(4):481–504
12. Amenta N, Choi S, Dey TK, Leekha N (2002) A simple algorithm for homeomorphic surface reconstruction. Int J Comput Geom Ap 12(1–2):125–141
13. Chang MC, Leymarie FF, Kimia BB (2009) Surface reconstruction from point clouds by transforming the medial scaffold. Comput Vis Image 113(11):1130–1146
14. Dey TK, Goswami S (2003) Tight cocone: a watertight surface reconstructor. In: the 8th ACM symposium on solid modeling and applications. New York, pp 127–134
15. Dey TK, Goswami S (2006) Provable surface reconstruction from noisy samples. Comp Geom 35(1–2):124–141
16. Dey TK, Giesen J, Hudson J (2001) Delaunay based shape reconstruction from large data. In: The IEEE symposium on parallel and large-data visualization and graphics. San Diego, CA, USA, pp 19–146
17. Amenta N, Choi S, Kolluri R (2001) The power crust. In: The 6th ACM symposium on solid modeling and applications, pp 249–260
18. OuYang D, Feng HY (2011) Reconstruction of 2D polygonal curves and 3D triangular surfaces via clustering of Delaunay circles/spheres. Comput Aided Design 43(8):839–847

The Analysis of Change Region About Networked Control System Based on the Behavior Profile

Xiangwei Liu, Mimi Wang and Lu Liu

Abstract In the analysis of Networked Control System, sometimes there may appear a problem of change, it is a problem that how to find the change regions. Previously mostly on the basis of determined change activities, to analyze direct behavior, to study change region. On the basis of analysis of direct and indirect constraints between behaviors, extending behavioral profiles into comprehensive behavioral profiles, a new method to locate dynamic change transitions is proposed, and on the basis of which analytical methods to determine the smaller change region is given in the paper. The theoretical analysis and specific example show that the method is very effective.

1 Introduction

Now, research on networked control system emerges as there requirements of actual control system development and draws more and more attention of the scientists and engineering technicians. For the same control system, analyst and modelers create different models, but not all models are to meet the requirements, which produced a problem of model change [1].

X. Liu (✉)
School of Management, Anhui University of Science and Technology, Huainan 232001, China
e-mail: lxw7710@tom.com

M. Wang · L. Liu
School of Science, Anhui University of Science and Technology, Huainan 232001, China
e-mail: weilu1100@163.com

L. Liu
e-mail: luliu@tom.com

For networked control system, there were some previous studies, but the majority is not on the basis of process behavior [2–6]. Van de Wouw et al. [2] developed a prescriptive framework for the stabilizing controller design based on approximate discrete-time models for nonlinear Networked Control Systems. Postoyan and Nesic [3] considered a general class of nonlinear reduced-order observers and showed that the global asymptotic convergence of the observation error in the absence of network-induced constraints is maintained for the emulated observer semi globally. Lopez-Echeverria and Magana [4] analyzed a new alternative approach to compensate for the effects of time delays on a dynamic network control system. Emeka et al. [5] presented the Network Control System Wind Tunnel, an integrated modeling and simulation tool for the evaluation. Zeng and Chow [6] addressed the issue of the tradeoff between NCS security and the NCS real-time performance. Previously, there are many scholars in the study of locating change. Guo et al. [7] modeled and analyzed some typical insulation defects. Dam and Winikoff [8] not only considered a number of changes, but also considered the model elements, which specified a consistent constraint using target constraint language and the same modeling language, and proposed a new way agent-oriented to repair this change. Fei et al. [9] analyzed the change communication between components based on a matrix method, and to find design conflicts of change in the Produce design phase. Eramo et al. [10] put forward a methods about change discovery of convergence of multiple architectural language, which neither rely on the model of synchronization nor depend on the model consistent.

The above method based on behavioral semantics change to found change, just were from the semantics point of view, did not real research to find the essence of the change from the perspective of the process behavior. At the moment, to study change from the perspective of the process behavior to study the change, is mainly from the behavior of similar measurements to detect change. However there is a certain deviation about the method discovering the model change, [11] put forward a method of RPST (Refined Process Structure Tree) which took a Single-Entry-Single-Exit structure into account. Weber et al. [12] proposed a method to determine change, by analyzing the relationship between before and after the move of the process model. Matthias [13] proposed a method based on the behavioral profiles to look for change, but which considered only the direct behavior relation and did not consider the uncertainty of change transitions.

Based on the above background, a method to look for dynamic change transition and to locate change regions in the networked control system is proposed.

2 Basic Concepts

This section describes the basic concepts and definitions used in our approach, notion of Petri net and weak order can reference literature [13], below we mainly introduce the concept of comprehensive behavior profiles.

Definition 1 (Comprehensive Behavioral Profile)Let (N, M_0) be a net system with $N = (P, T, F)$. A pair of transitions $(x, y) \in (T \times T), (y, z) \in (T \times T)$, can be in the following profile relations:

1. The simulated strict order relation $x \overset{\rightarrow}{} z$, iff $x \succ y$, $y \nsucc x$, $y \succ z$ and $z \nsucc y$.
2. The simulated exclusiveness relation $x \overset{-}{+} z$, iff $x \nsucc y$, $y \nsucc x$, $y \succ z$ and $z \nsucc y$.
3. The simulated interleaving relation $x \overset{-}{\|} z$, iff $x \succ y$, $y \succ x$, $y \succ z$ and $z \nsucc y$.

$\quad$ B $= \left\{ \rightarrow, +, \|, \overset{\rightarrow}{}, \overset{-}{+}, \overset{=}{\|} \right\}$ is the comprehensive behavioral profile over T.

$\quad$ Before looking for change region, we must first determine the specific location of the change transition, first give the definition of change transition as follows.

Definition 2 [13] Let $S_1 = (N_1, M_1)$ and $S_2 = (N_2, M_2)$ be net systems with $N_1 = (P_1, T_1, F_1)$ and $N_2 = (P_2, T_2, F_2), B_1 = \{\overset{\rightarrow}{}_1, +_1, \|_1, \overset{\rightarrow}{}_1, \overset{-}{+}_1, \overset{=}{\|}_1\}$ and $B_2 = \{\overset{\rightarrow}{}_2, +_2, \|_2, \overset{\rightarrow}{}_2, \overset{-}{+}_2, \overset{=}{\|}_2\}$ their behavioral profiles, and $\sim \subseteq T_1 \times T_2$ a comprehensive behavioral profile consistent correspondence relation. Let t_x be a change transition of a consistent change in S_1, it holds that:

1. $P_{C_1} = T_2 \backslash PT \backslash ST \backslash BT$
2. $P_{C_2} = T_2 \backslash EIT \backslash EIT^- \backslash IIT \backslash IIT^- \backslash ET \backslash IT$

Definition 3 (Suspected Change Region) A suspected change region C_r is a region where t_x is the center, and the region is surrounded by all transitions which are in direct or indirect behavior relation with t_x.

Definition 4 (Change Region) A change region C_P is a subset of anyone suspected change region, and $C_P = C_{r_1} \cap C_{r_2} \cap C_{r_3} \cdots \cap C_{r_n}$.

$\quad$ Other definitions can see [13].

3 The Analysis Method of Looking for Change Region Based on the Dynamic Behavior

When looking for the change transitions, because do not sure the specific location of the change transition, it is cumbersome and error-prone if checking up transitions one by one, so the sake of simplicity, we can take the regional block exclusion, fist analyze anyone transition which is regarded as a suspected change transition within the region. Following we introduce the concept of relation circle of suspected change transition.

Definition 5 (Relation Circle of Suspected Change Transition) A relation circle Θ_t of a suspected change transition t_x, is a circle which makes t_x as the center,

1. If $P_{C_1} \neq \varnothing \wedge P_{C_2} \neq \varnothing$, it is a square region which makes $m = \min\{\|P_{C_1}\|, \|P_{C_2}\|\}$ as it's a side, with $\|P_{C_1}\|, \|P_{C_2}\|$ show the number of transitions of simulated boundary transition reduction and simulated inter-boundary transition reduction, respectively.
2. If $P_{C_1} = \varnothing \wedge P_{C_2} \neq \varnothing$, it is a square region which makes $m = \max\{\|P_{C_1}\|, \|P_{C_2}\|\}$ as it's a side (similarly, we can see the condition of $P_{C_1} \neq \varnothing \wedge P_{C_2} = \varnothing$).
3. If $P_{C_1} = \varnothing \wedge P_{C_2} = \varnothing$, it is a region surrounded by all transitions which are in direct or indirect behavior relation with t_x. Follow, we can look for the change transitions.

We can see that: (1) the more suspected change regions, the smaller its intersection, then the smaller the change region we found, so in real life, we should find these suspected change regions as much as possible, to obtain much smaller change region, reducing the trouble for handling fault; (2) After found C_{r_1}, because of C_P is a subset of C_{r_1}, the next step when looking for the next suspected change transition, we can check out the transitions in C_{r_1};, (3) to be less onerous, we see that it is smaller the region C_{P_1} obtained if P_{C_1} and P_{C_2} produced by the suspected change transition are non-empty sets, than the region C_{P_2} obtained if P_{C_1} and P_{C_2} produced by the suspected change transition are at least one empty set, that is $C_{P_1} \subset C_{P_2}$, then it holds that $C_{P_1} \cap C_{P_2} = C_{P_1}$;, (4) if P_{C_1} and P_{C_2} produced by the suspected change transition are empty sets, which is different with above conditions, such as if P_{C_1} and P_{C_2} produced by the suspected change transition t_1, and P'_{C_1}, P'_{C_2} produced by the suspected change transition t'_1 are non-empty sets, also $P'_{C_1} \supseteq P_{C_1}, P'_{C_2} \supseteq P_{C_2}$, obtaining change region C_{P_1}, C'_{P_1} respectively, then $C'_{P_1} \subseteq C_{P_1}$, and $C'_{P_1} \cap C_{P_1} = C'_{P_1}$. So when looking for the change region, we should look for the suspected change transition which obtains the more transitions in P_{C_1}.

Algorithm 1: locate the change regions

(1) For anyone transition t_x, according to definition 2, definition 2, work out P_{C_1}, P_{C_2} respectively, if $P_{C_1} = \varnothing \wedge P_{C_2} = \varnothing$, then t_x is not a change transition, return to step (2); otherwise, t_x is a change transition, then return to step (7).
(2) According to definition 5, work out relation circle of suspected change transition Θ_{t_1} generated by t_x, record the set of transitions in Θ_{t_1} as T_{11}, return to step (3).
(3) Analyze $t_y \in T_1 \backslash T_{11}$, according to definition 2, work out P_{C_1}, P_{C_2} respectively, if $P_{C_1} = \varnothing \wedge P_{C_2} = \varnothing$, then t_y is not a change transition, return to step (4); otherwise, t_y is a change transition, then return to step (7).
(4) According to definition 5, work out relation circle of suspected change transition Θ_{t_2} generated by t_y, record the set of transitions in Θ_{t_2} as T_{12}, return to step (5).
(5) Analyze $t_z \in T_1 \backslash T_{11} \backslash T_{12}$, according to definition 2, work out P_{C_1}, P_{C_2} respectively, if $P_{C_1} = \varnothing \wedge P_{C_2} = \varnothing$, then t_z is not a change transition, return to step (6); otherwise, t_z is a change transition, then return to step (7).

(6) Similarly, implement above steps, until some transition $t_m \in$ $T_1 \backslash T_{11} \backslash T_{12} \backslash \cdots \backslash T_{1m}$, it holds $P_{C_1} \cup P_{C_2} \neq \varnothing$, then t_m is a change transition, then return to step (7).

(7) Analyze the change transition t, if $P_{C_1} \neq \varnothing \wedge P_{C_2} \neq \varnothing$, according to definition 3, work out the suspected change region C_{r_1}, then return to step (8); otherwise, if $P_{C_1} \neq \varnothing \wedge P_{C_2} = \varnothing$ (the similar as $P_{C_1} = \varnothing \wedge P_{C_2} \neq \varnothing$), analyze the transition in the P_{C_1} (similarly analyze the transition in the P_{C_2}), return to step (8).

(8) Analyze the boundary transition t_1 in C_{r_1}, according to definition 2, work out P_{C_1}, P_{C_2} respectively, if $P_{C_1} \neq \varnothing \wedge P_{C_2} \neq \varnothing$, then t_1 is a change transition, according to definition 4, work out the suspected change region C_{r_2}, then return to step (9); otherwise, if $P_{C_1} = \varnothing \wedge P_{C_2} = \varnothing$, then t_1 is not a change transition, then $C_{r_2} = C_{r_1} \backslash \{t_1\}$, return to step (9).

(9) Analyze the another boundary transition t_2 in C_{r_1}, according to definition 2, work out P_{C_1}, P_{C_2} respectively, if $P_{C_1} \neq \varnothing \wedge P_{C_2} \neq \varnothing$, then t_2 is a change transition, according to definition 3, work out the suspected change region C_{r_3}, then return to step (10); otherwise, if $P_{C_1} = \varnothing \wedge P_{C_2} = \varnothing$, then t_2 is not a change transition, then $C_{r_3} = C_{r_2} \backslash \{t_2\}$, return to step (10).

(10) Similarly, implement above steps, until the last boundary transition t_i in C_{r_1}, according to definition 2, work out P_{C_1}, P_{C_2} respectively, if $P_{C_1} \neq \varnothing \wedge P_{C_2} \neq \varnothing$, then t_i is a change transition, according to definition 3, work out the suspected change region C_{r_i}, algorithm terminates; otherwise, output $C_P = C_{r_1} \cap C_{r_2} \cap C_{r_3} \cdots \cap C_{r_i}$; otherwise, if $P_{C_1} = \varnothing \wedge P_{C_2} = \varnothing$, then t_i is not a change transition, then $C_{r_i} = C_{r_i} \backslash \{t_i\}$, algorithm terminates, output $C_P = C_{r_1} \cap C_{r_2} \cap C_{r_3} \cdots \cap C_{r_i}$.

4 Case Study About the Networked Control System

First, the Petri net of a networked control system model and target model are shown in Figs. 1 and 2.

For any suspicious change transitions in Fig. 2, for example t_{16}, it is in the direct behavior relation with $t_{11}, t_{17}, t_{15}, t_{21}$, and it is in indirect behavior relation with $t_{10}, t_7, t_9, t_{18}, t_{19}$ which holds that $PT = \{t_{11}, t_{14}\}, ST = \{t_{21}, t_{18}\}, BT = \{t_{17}, t_{15}, t_{19}\}$, so $P_{C_1}(t_{16}) = T_2 \backslash PT \backslash ST \backslash BT = T_2 \backslash \{t_{14}, t_{11}, t_{17}, t_{15}, t_{19}, t_{21}, t_{18}\}$, $EIT = \varnothing, EIT^- = \varnothing, IIT = \{t_{15}\}$, $IIT^- = \{t_{19}\}, ET = \varnothing$, $P_{C_2}(t_{16}) = T_2 \backslash EIT \backslash EIT^- \backslash IIT \backslash IIT^- \backslash ET \backslash IT = T_2 \backslash \{t_{15}, t_{19}\}$, and P_{C_1}, P_{C_2} are not non-empty sets, so t_{16} is the change transition, and change region found from change transition t_{16} that composed of $t_{10}, t_{11}, t_{17}, t_{16}, t_{15}, t_{19}, t_{21}$ in Fig. 2 is the first suspected change region (like Fig. 3a). Below to find the second suspected change transition

Fig. 1 Petri net model of a networked control system
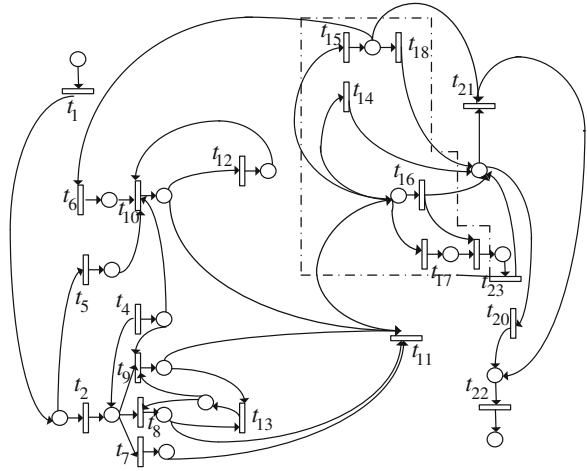
Fig. 2 Petri net of target model
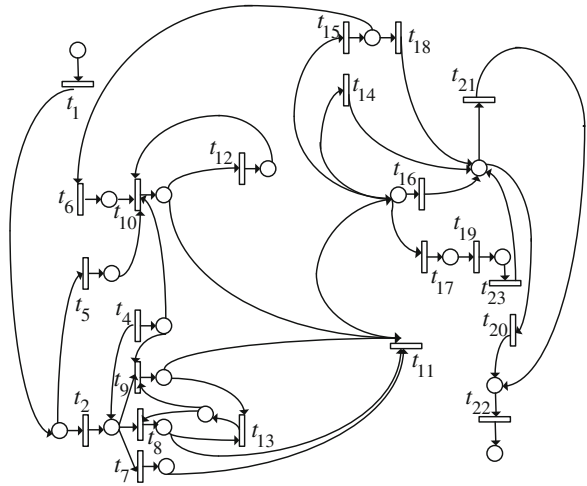
in accordance with conditions (2) of the third section, that is, the change region C_{r_1} in Fig. 3, we can found that t_{21}, t_{10} are not change transitions, and for the suspected change transition t_{15}, t_{17}, obtain C_{r_2}, C_{r_3} there, it holds that $C_{r_3} = C_{r_2} = C_{r_1}$. In the same way, t_{21} regards as the suspected change transition, C_{r_4} is surrounded by $t_{11}, t_{17}, t_{16}, t_{15}, t_{19}, t_{21}$ (like Fig. 3b), $C_{r_4} \cap C_{r_3} \cap C_{r_2} \cap C_{r_1} = C_{r_4}$; when t_{21}, t_{11} regards as the suspected change transition, it holds C_{r_5}, C_{r_6} (like Fig. 3c). $C_{r_5} \cap C_{r_4} \cap C_{r_3} \cap C_{r_2} \cap C_{r_1} = C_{r_4}.C_{r_6}$ is surrounded by $t_{11}, t_{17}, t_{16}, t_{15}, t_{19}$, then $C_{r_6} \cap C_{r_5} \cap C_{r_4} \cap C_{r_3} \cap C_{r_2} \cap C_{r_1} = C_{r_6}$. So that change region is surrounded by $t_{11}, t_{17}, t_{16}, t_{15}, t_{19}, t_{14}, t_{18}$, i.e. the dotted line in Fig. 1.
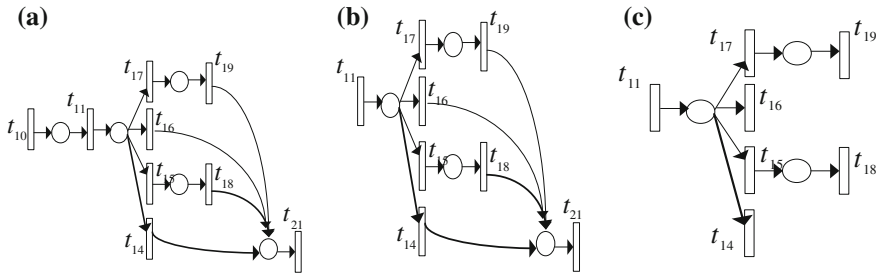
Fig. 3 Suspicious change regions

5 Conclusion

In this paper, on the basis of previous study, we extend the behavioral profile, define the three indirect behavior relations on the basis of the weak order relation, extend the behavior profile into a comprehensive behavioral profile, and propose a method to look for dynamic change transition and so as to locate change regions, last give the analysis for an example about a networked control system.

Still, we also have to reflect on some limitations of our approach. The approach takes into account the behavior relationship based on comprehensive behavioral profile to look for change regions, but not the marking, for some complexity situations, it will be the future work.

References

1. Christian G, Jochen MK, Gregor E (2009) Language-independent change management of process models. Lect Notes Comput Sci. doi:10.1007/978-3-642-04425-0_12
2. Van de Wouw N, Nesic D, Heemels WPMH (2012) A discrete-time framework for stability analysis of nonlinear networked control systems. Automatica 48:1144–1151
3. Postoyan R, Nesic D (2012) On emulated nonlinear reduced-order observers for networked control systems. Automatica 48:645–652
4. Lopez-Echeverria D, Magana ME (2012) Variable sampling approach to mitigate instability in networked control systems with delays. IEEE Trans Neural Networks Learn Syst 23:119–126
5. Emeka E, Jia B, Riley D, Weng JN, Wei Y, Xue Y, Koutsoukos X, Sztipanovits J (2012) NCSWT: an integrated modeling and simulation tool for networked control systems. Simul Model Pract Theor 27:90–111

6. Zeng W, Chow MY (2012) Optimal tradeoff between performance and security in networked control systems based on coevolutionary algorithms. IEEE Trans Industr Electron 59:3016–3025
7. Guo CX, Huang CJ, Qian Y, Liu JH, Jiang XC (2009) The electro quasi static field simulation and analysis for the insulation defects in XLPE power cables. Int Rev Elect Eng IREE 4:1413–1417
8. Dam HK, Winikoff M (2011) An agent-oriented approach to change propagation in software maintenance. Auton Agent Multi-Agent Syst 23:384–452
9. Fei GY, Gao J, Owodunni O, Tang XQ (2011) A method for engineering design change analysis using system modeling and knowledge management techniques. Comput Integr 24:535–551
10. Eramo R, Malavolta I, Muccini H, Pierantonio A (2012) A model-driven approach to automate the propagation of changes among architecture description languages. Softw Syst Model 11:29–53
11. Vanhatalo J, Völzer H, Koehler J (2009) The refined process structure tree. Data Knowl Eng 68:793–818
12. Weber B, Reichert M, Rinderle-Ma S (2008) Change patterns and change support features—enhancing flexibility in process-aware information systems. Data Knowl Eng 66:438–466
13. Matthias W (2011) Behavioural profiles—a relational approach to behaviour consistency. Institutional Repository   of the University of Potsdam. http://opus.kobv.de/ubp/volltexte/2011/5559/URNurn:nbn:de:kobv:517-opus-55590(2011)

The Application of Ant Colony Optimization in CBR

Jianhua Shu

Abstract In the Case-Based Reasoning (CBR) System, the retrieval efficiency and system performance are reduced because of the unlimited increasing case base with the incremental learning. This paper proposes the method of ant colony optimization (ACO) in the CBR system. This method combines the increased efficiency of case retrieval, the effective case base indexing, and the validity of maintenances by adding or reducing cases. Through the all processes we have used the clustering and classification algorithm based ACO. The implementation of the ACO algorithm into the CBR system is successful and the experimental results verify its effectiveness.

Keywords CBR · ACO · Clustering · Classification · Case retrieval · Case-base maintenance

1 Introduction

Case-Based Reasoning (CBR) solves new problems by reusing the old solutions in similar context. As an effective problem-solving method, CBR has been deployed in a wide variety of applications, such as planning, classification, diagnosis, decision supporting system and others [1].

CBR as an important complement to the rule-based reasoning (RBR), to some extent, has make up for the defects in the RBR. The knowledge acquisition

J. Shu (✉)
School of Medicine Information Engineering, Anhui University
of Traditional Chinese Medicine, Hefei, Anhui, China
e-mail: gaiersitu@gmail.com

J. Shu
School of Mathematics and Physics, Anhui University of Science
and Technology, Huainan, Anhui, China

bottleneck has received the attention of artificial intelligence researchers. CBR has many advantages which are discussed below: it has a strong learning ability; the realization is relatively simple; the maintenance of knowledge is simple. However, CBR, there are also some problems: (1) it is particularly sensitive to noise data, error data and redundant data which will affect the efficiency of the system retrieves and solution results easily; (2) when the number of cases in the case base is increasing, it may be retrieved low efficiency, redundant search results; (3) the knowledge acquisition bottlenecks still exist. Therefore, it is necessary to improve the search capabilities of the case base and strengthen the case base maintenance.

Ant colony optimization (ACO) [2] is a population-based meta heuristic that can be used to find approximate solutions to difficult optimization problems. It is inspired by the foraging behavior of ant colonies. The first ant colony optimization algorithm was proposed in 1991 by Dorigo M for the approximate solution of the traveling salesman problem. Afterwards, the ACO algorithm has been applied successfully to many combinatorial optimization problems, such as the assignment problems, scheduling problems, vehicle routing problems. But, recently there exist new ACO algorithms developed for clustering and classification problems used also in data mining [3, 4].

The aim of this study is solving the application of clustering and classification problems based ACO in CBR. In the CBR system, the case base is becoming increasingly larger with the incremental learning which results in the decline of case retrieval efficiency and its weaker performance. This paper proposes the ACO algorithm using in the retrieval of the case, the indexing of the case and the maintenance of the case base to solve those problems when the case library is of massive size and the decline of efficiency occurs. Finally the implementation of our proposed method into the CBR system is successful and the experimental results verify its effectiveness.

2 The Process of CBR and Case-Base Maintenance

CBR solves new problems by reapplying and adapting solutions that were used to solve old similar problems. A typical CBR process is comprised of four-Re steps: (1) Retrieve: given a new problem, retrieve the most similar cases in the library; (2) Reuse: apply the solutions from these retrieved cases to the new problem; (3) Revise: use the retrieved solution and adapt it as needed to solve the new problem; (4) Retain: the new achieved solution can be added to the library as a new case.

The CBR is an incremental learning, so that CBR can solve more problems. Over time, new cases are gradually added, and the case base is growing, so that some problems such as beyond the design capacity of the case base that limit the size of the case base, case retrieval efficiency reduce, knowledge redundancy, are appeared. Therefore, the CBM (Case Base Maintenance) is the most important process in the CBR system. CBM is mainly based on the feedback from the users.

If the user is satisfied, and the case was set for each search really solved the problem temporarily, you do not need to make changes on the case; if retrieved case can not meet the users' needs, the maintenance personnel need to revise the case in the case base based on the user's feedback and combined with the expert knowledge. If the new problem can not be retrieved by using the case base, we need to add the new case to the case base. If the case base has some redundant information, the retrieval speed of the target case is too slow, or the search results have the redundant information, we need to reduce the redundant case information according to the actual situation. If the cases are obsolete, we need to revise some of the information in the case base. CBM is the most important step in the safe and reliable CBR system. The inappropriate operations of CBM will lead to the lack and the redundancy of the knowledge in the CBR system, the reduced retrieval efficiency [5].

3 The Application of ACO Algorithm in CBR

In this section, the ACO algorithm using in CBR method is described in detail. This method consists of the clustering based on ACO and the classification based on ACO, and the process using them in the CBR system. The process includes the retrieval of the case, the creation of the case base, the indexing of the case and the maintenance of the case base. The details of the improved ACO-based clustering algorithm and the improved ACO-based classification algorithm and the process of using them in CBR are described in following sub-sections.

3.1 Improved ACO-Based Clustering Algorithm

The paper [6] proposes an improvement of clustering algorithm based on ACO, to settle the defect of low efficiency for the lack of pheromone in the initial stage. To solve the problems of the existed clustering algorithm, we improved the ACO algorithm by amending the distribution of the initial pheromone, the generation of the initial cluster set and the termination conditions of the algorithm, and so on. This method increased the utilization of the pheromone to assist the generation of the cluster set.

The general mathematical description: to let the data set for $X = \{X_i, i = 1, 2, \ldots, n\}$, where X_i is the p-dimensional vector, the clustering is to find a partition of $C = (C_1, C_2, \ldots, C_m)$ by the conditions as follows,

$$X = \bigcup_{i=1}^{m} C_i, \quad C_i \neq \Phi(i = 1, 2, \ldots, m) \tag{1}$$

$$C_i \cap C_j = \Phi, \quad (i, j = 1, 2, \ldots, m; i \neq j). \tag{2}$$

And the sum of the dispersion given by

$$J = \sum_{k=1}^{m} \sum_{X_i \in C_k} d(X_i, Z_k). \tag{3}$$

When J reached the minimum value, where Z_k is the center of the kth cluster and $d(X_i, Z_k)$ is the distance of the sample to the cluster center.

The basic procedure of the improved ACO-based clustering algorithm is described as follows:

Step 1: Input data set; Initialize the pheromone Matrix, $\tau(r,s) = 1/d(r,s)$, where $\tau(r,s)$ is the pheromone concentration of the path, and $d(r,s)(r,s = 1, 2, \ldots, n)$ is the distance of the path;

Step 2: Generate initial cluster sets by the pheromone concentration;

Step 3: Set the pheromone distribution among the initial cluster in the initial cluster;

Step 4: Place m ants randomly to the initial cluster sets;

Step 5: Choose the j ant to tournament selection mechanism to visit the initial cluster sets in the result sets; do until new results are generated;

Step 6: Compute the sum of the dispersion, save the best result;

Step 7: update the pheromone concentration of initial cluster sets;

Step 8: Computing the sum of the dispersion $J = \sum_{k=1}^{m} \left[\frac{1}{s_k} \sum_{X_i \in C_k, X_j \in C_k} d(X_i, X_j) \right]$, $(i \neq j, s_k in C_k)$, if the result is the need, output the best clustering results, else, select next undone search go to step 4.

3.2 Improved ACO-Based Classification Algorithm

There is a new method based ACO of mining classification rule from categorical database [7]. Firstly, to improve the processing capability of the algorithm on large data sets, we use ACO into attributes reduction for minimizing dataset. Then, we use ACO for discovering classification rule. This algorithm follows a sequential covering approach to discover a list of classification rules covering all, or almost all, the training cases. At first, the list of discovered rules is empty and the training set consists of all the training cases.

Each cycle consists of three steps, comprising rule construction, rule pruning, and pheromone updating, discovers one classification rule. This rule is added to the list of discovered rules and the training cases that are covered correctly by this rule (i.e., cases satisfying the rule antecedent and having the class predicted by the rule consequent) are removed from the training set. This process is performed iteratively while the number of uncovered training cases is greater than a user-specified threshold, called Max_uncovered_cases. Rule pruning is a commonplace technique, so we redefine a new pruning procedure.

The basic procedure of the improved ACO-based classification algorithm is described as follows:

Step 1: Simplify data sets /* attributes reduction */
Step 2: Split data sets /* divide into the test set and the training set */
Step 3: result rule sets = [] /* the rule sets is initialized to empty */
Step 4: Does a cycle have three steps: comprising rule construction, rule pruning, and pheromone updating
Step 5: choose the best rule R_{best} from R_i
Step 6: add R_{best} into the result sets, reduce the training sets
Step 7: do until the number of uncovered training cases is greater than Max_ uncovered_cases, output the result rule sets.

3.3 ACO and Data Mining in CBR

In this sub-section, the ACO algorithm using in CBR method is described in detail. The framework of the ACO in CBR is shown in Fig. 1.

In the CBR system, the case base is becoming increasingly larger with the incremental learning which results in the decline of case retrieval efficiency and its weaker performance. So, we must solve the following problems: (a) To increase the efficiency of the case retrieval; (b) To establish the effective case base indexing; (c) do some valid maintenances by adding or reducing. The details of these are described as follow.
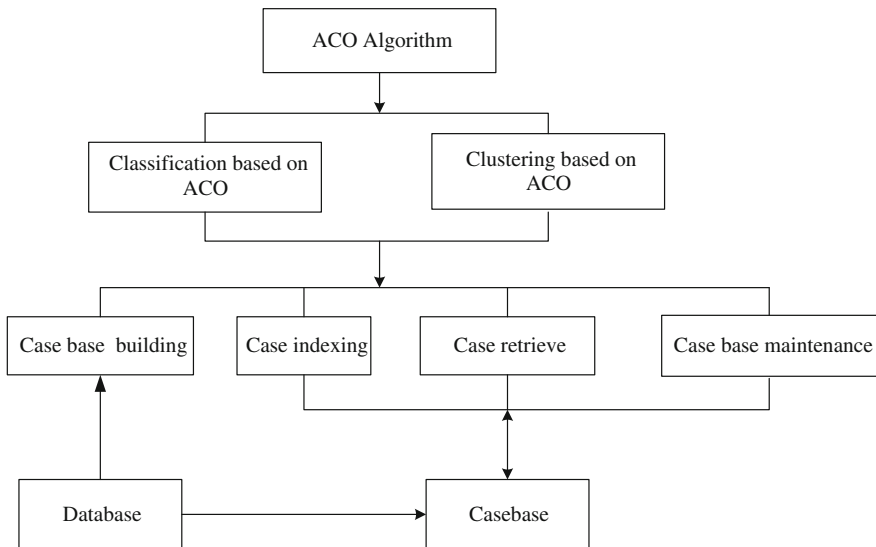

Fig. 1 The framework of the ACO in CBR

The case retrieval is to search for one or more of the most similar case with the current problem in the case base. The case retrieval needs to achieve the following two objectives: (a) The retrieved cases are as little as possible; (b) The retrieved cases are related or similar as possible with the target case. Clustering algorithm has been used in CBR [8], so we use the improved ACO-based clustering algorithm as the method of case retrieval.

In the process of case retrieval, we choose the cluster set combined with the nearest neighbor retrieval algorithm. That is, first we must determine which cluster set the target case belongs, and then we use the nearest neighbor retrieval algorithm to find the most similar to the target case in the cluster set.

To create the effective case base indexing, we use the improved ACO-based clustering algorithm to generate the cluster sets. We make the center of cluster results set and the cases of an independent cluster set called class 0 as the index of case base. If the similarity of the case and any center of cluster set is very low (we can set a threshold), we would make them into class 0 instead of dropping them.

In the CBM sub system, we use the follow strategies: (a) when we can not find the suitable old case to solve the new problem, the experts would be invited. The solutions should be added as new cases into case base; (b) we use the improved ACO-based clustering algorithm to generate the cluster sets. Then, we compute the reach ability set of each case in a cluster, if the reach ability set of a case is large then we would delete that case and if some cases have same reach ability set we would delete the case whose coverage set is smaller.

Similarly, the improved ACO-based classification algorithm can also use in the CBR system [9]. This algorithm can be used to generate rules. We match a new case to the rules. If the only rule is to match, the new case is classified to the class represented by the rule. When no rules match with the new case, we would complete a partial match to determine whether some attribute values are matching. To order the cases in the case base by using k-NN algorithm we can get the similar set which has k cases. If all the cases in the similar set can predict to class C, the new case would be added into the class C. We can delete the redundancy or error cases obtained by the k-NN algorithm.

4 Experimental Method and Results

To test and verify the effectiveness of the ACO in CBR, on the basis of the analysis in the previous algorithms, we developed the experimental platform use of Microsoft Visual Studio. Net 2003 (C#). The performance was evaluated using four public-domain data sets from the University of California at Irvine repository.

The main characteristics of the data sets used in our experiment are summarized in Table 1. The first column gives the data set name, while the other columns indicate, respectively, the number of cases, the number of attributes, and the number of classes of the data set.

Table 1 Original data set of experiment

Data sets	#Cases	#Attributes	#Classes
Tic-tac-toe	958	9	2
Breast cancer	683	9	2
Soybean(I)	47	35	4
Soybean(II)	376	35	19

We mainly analyze the conclusions of the ACO-based classification algorithm and the ACO-base clustering algorithm in case retrieval [6, 7]. The ACO-base clustering algorithm has used in the CBR system. For Small Soybean data set of 47 cases, the correct retrieval rate is 100 %. For the Soybean data set of 376 test cases, the system correctly retrieved 343 cases, the system retrieved correctly rate of 91.22 %. To use the ACO-based classification algorithm, Cross-validation 10 times for Tic-tac-toe data sets, the prediction accuracy rate is 78.89+/−3.23 %. Cross-validation 10 times for Breast cancer data set, the prediction accuracy rate is 79.59+/−2.34 %. Therefore, the ant colony algorithm in data mining research CBR system can effectively improve the performance of case system.

We can further verify the efficiency of the ACO-based classification algorithm and the ACO-based clustering algorithm in the maintenances of the case base by the average similarity. That is, when we add or reduce some cases in the case base, by comparing the average similarity of the test cases which are according to the operations we take before and after. The high average similarity shows the method's effective. We can also compare the retrieval rate to verify the efficiency.

5 Conclusions

In the CBR system, the case base is becoming increasingly larger with the incremental learning which results in the decline of case retrieval efficiency and its weaker performance. In this paper, a new and efficient data mining method, called clustering and classification method based on ACO, has been proposed for the application in the CBR system. We mainly solve the following problems: (a) To increase the efficiency of the case retrieval; (b) To establish the effective case base indexing; (c) Do some valid maintenances by adding or reducing. Two important directions for future research are as follows. First, it would be interesting to introduce other intelligence algorithm such as PSO (particle swarm optimization), DNA computing into CBR system. Second, it would be interesting to use these algorithms to improve the efficiency of knowledge acquisition in the knowledge management system.

References

1. Shanlin Y, Zhiwei N (2004) Machine learning and intelligent decision support system. Science Press, Beijing, pp 79–80
2. Dorigo M, Blum C (2005) Ant colony optimization theory: a survey. Theoret Comput Sci 344:243–278
3. Parpinelli RS, Lopes HS, Freitas A (2002) A data mining with an ant colony optimization algorithm. IEEE Trans Evol Comput 6(4):321–332
4. Shelokar PS, Jayaraman VK, Kulkarni BD (2004) An ant colony approach for clustering. Anal Chim Acta 509:187–195
5. Zehraoui F, Kanawati R, Salotti S (2003) Case base maintenance for improving prediction quality. In: Ashley KD, Bridge DG (eds) ICCBR 2003, LNAI 2689. Springer, Heidelberg, pp 703–717
6. Ni Z, Shu J, Xu L (2007) The case retrieval method with an ant colony optimization algorithm. Appl Res Comput 24(S):1289–1290
7. Shu J, Ni Z, Yang S (2007) A mining classification-rule method based on an ant colony optimization algorithm. J Guangxi Normal Univ (Nat Sci Ed) 25(4):18–23
8. Qiao L, Jiang HL, Jia SJ (2011) Case retrieval strategy based on improved k-means clustering. Comput Eng 37(5):193–195
9. Khelassi A (2012) Data mining application with case based reasoning classifier for breast cancer decision support. In: Proceedings of MICIT, Liverpool, UK, pp 1–6

Genetic Algorithm for Solving Survivable Network Design with Simultaneous Unicast and Anycast Flows

Huynh Thi Thanh Binh, Son Hong Ngo and Dat Ngoc Nguyen

Abstract We consider the survivable network design problem for simultaneous unicast and anycast flow requests. In this problem, a network is modeled by a connected, weighted and undirected graph with link cost follows All Capacities Modular Cost (ACMC) model. Given a set of flow demand, this problem aims at finding a set of connection with minimized network cost to protect the network against any single failure. This problem is proved to be NP-hard. In this paper, we propose a new Genetic Algorithm for solving the ACMC Survivable Network Design Problem (A-SNDP). Extensive simulation results on Polska, Germany and Atlanta network instances show that the proposed algorithm is much more efficient than the Tabu Search and other baseline algorithms such as FBB1 and FBB2 in terms of minimizing the network cost.

Keywords Survivable network design · Anycast · Unicast · Genetic algorithm

1 Introduction

There are many types of connection for data transmission over the Internet. The two most popular types of connection are unicast and anycast. A connection from a node to another is called unicast. An anycast is also a connection from a node to

H. T. T. Binh (✉) · S. H. Ngo · D. N. Nguyen
School of Information and Communication Technology, Hanoi University of Science
and Technology, Hanoi, Vietnam
e-mail: binhht@soict.hut.edu.vn

S. H. Ngo
e-mail: sonnh@soict.hut.edu.vn

D. N. Nguyen
e-mail: dat.thientai@gmail.com

another, the difference is that the destination node has a one or many replica servers which back up for it. Anycast connection is currently used in many applications such as Domain Name Service (DNS), Web Service, Overlay Network, peer-to-peer (P2P) systems, Content Delivery Network (CDN), software distribution [1]. The popularity of anycast technology will increase in the near future, since many new services that use both unicast and anycast paradigms are being developed [2].

In the Internet, any network failure can cause serious consequences, e.g., a single link failure affected more than 30,000 users and it takes 12 h to fix [3]. Therefore, the design of survivable networks is a crucial problem. In the survivable network design problem (SNDP) with simultaneous unicast and anycast flows, the common objective is to minimize the network cost to protect network against failures. To guarantee the survivability, we adopt the protection approach [1, 4–6] in that each connection must include a working path and a link-disjoint backup path. The working path is used for data transmission in failure-free state of the network. If there is a single link failure on the working path, the failed connection is then switched to the backup path.

In [1], Gładysz et al. have considered the SNDP for networks using ACMC (All Capacities Modular Cost) link cost model (here after called by A-SNDP). In the ACMC model, a link has many bandwidth levels, each level has a corresponding cost. Many demands can go over the same link thus the link cost is defined as the total of required bandwidth from all demands on that link. The network cost is defined as the total of all link cost.

The A-SNDP problem is defined as follows. Given an undirected graph $G = (V, E)$ and a set of customer demand between node pairs with corresponding bandwidth and type of connection (anycast or unicast). The goal is to find a set of connection for all demands such that the network cost (*NCost*) is minimized:

$$NCost = \sum_i c_i$$

where $c_i = C_k$, if $B_{k-1} < \sum R_{ij} < B_k$.

Here, c_i is cost of link i; B_k and C_k is bandwidth and corresponding cost in level k; R_{ij} is the required bandwidth from the demand j on the link i.

The authors in [1] have proposed a heuristic for A-SNDP using Tabu Search, however, their result is still far from optimal approach. In order to further minimize the network cost using heuristic approach, we have developed a new genetic algorithm for solving A-SNDP. In this paper, we first propose new individual encoding scheme called Connection Database based Encoding (CDE). We then design the evolution operators using CDE, simulate our proposed algorithm with three typical network instances [7] and compare the results with other approaches such as Tabu Search [1], FBB1 and FBB2 [8]. Results obtained in experimental works show that our proposed algorithm are better than the others in terms of minimizing the network cost.

The rest of this paper is organized as follows. Section 2 describes the related works. Our new approach to encode individual is shown in Sect. 3. Section 3.6 presents the proposed GA algorithm to solve A-SNDP. Experimental results are given in Sect. 4. The paper concludes with Sect. 5 by discussion and future works.

2 Related Works

The SNDP is generally presented in [9], when considering both economics and reliability in telecommunication network design. The SNDP has to guarantee the survivability of network system and also to minimize the network cost. The most popular way mentioned in many researches is the single backup path approach. The main idea of this method is as following: Each connection has a working path and a backup path, if there is a single link failure on the working path, the failed connection is then switched to the backup path [1, 4–6].

In the literature, there are many papers on minimizing the network cost for SNDP problem (see [1, 9, 10] and references there in). They use—and—bounds or branch—and—cut methods to find optimal solution. These methods can only use for networks with small number of nodes. For larger networks, they may propose heuristic such as evolutionary algorithms, tabu search [1] and simulated annealing [10]. In [9], Vissen and Gold apply the evolution strategy (ES). It is shown by the authors that when using ES, a larger population can help to achieve a better result than a smaller one by avoiding or delaying convergence on local suboptimal. However, this algorithm is applied in the network which has only unicast flows.

With the network which has both anycast and unicast flows, Walkowiak et al. presented a heuristic algorithm for solving A-SNDP [7]. The main idea of this algorithm is based on Flow Deviation method [10] and Local Search algorithm [11]. They achieve the quite good result with Polska (12 nodes, 36 links, 65 unicast, 12 anycast) network, the detail shows that the average gap of the proposed heuristic to optimal result is 7.11 %. Furthermore, they also built Tabu Search algorithm based on hill climbing algorithm with some heuristics to solve this problem [1]. Experiments on three large instances which are Polska (12 nodes, 36 links, 65 unicast, 12 anycast), Germany (17 nodes, 52 links, 119 unicast, 13 anycast), Atlanta (26 nodes, 82 links, 234 unicast, 22 anycast) showed many promising results. In particular with Polska network, they achieve the average gap to optimal results is 2.57 % for 70 % anycast/ 30 % unicast case and 2.00 % for 80 % anycast/ 20 % unicast case. However, their Tabu Search algorithm is still simple and their results cannot be optimal completely.

In [8], Huynh et al. also proposed two heuristics called FBB1 and FBB2 for solving A-SNDP. The main idea of FBB1 is based on utilizing the redundant bandwidth corresponding with paid cost level in each link. FBB2 is the combination of FBB1 and Tabu Search algorithm. Experiments on three network instances which are Polska, Germany and Atlanta network [1, 3–7, 9, 10, 12–15]. With each instance, 10 test sets are randomly created. The results show that their

proposed approach is quite effective with A-SNDP. On all instances, FBB1 and FBB2 have better results than Tabu Search in most of test sets. However, the gap between their results is still not too significant.

In the next sections, we introduce a new scheme to encode a graph that will be used in our new genetic algorithm for solving A-SNDP. We hope that the result found by proposed genetic algorithm will be further improved, especially on large size network instances.

3 Connection Database Based Encoding

3.1 Definition

A solution of A-SNDP is a set of connections that satisfy all customer's demand. To facilitate the evolution operators in GA, we propose a new scheme to pre-process a graph by encoding the candidate connections between two nodes. We call it by Connection Database based Encoding (CDE). A connection database corresponds to each demand. The connection database is built as following by a routine called FindOtherWays. The main idea of FindOtherWays is as following: First, we initialize an empty database with two columns: ID and Connection. For each connection demand from A to B, we use Shortest Path algorithm to find a connection between them and insert it into the column Connection of this database. The running time of Shortest Path is $O(n^2)$ for each demand, where n is the number of vertices of input graph. The ID value of this connection is indexed from 1. Next, we choose a random link in this connection and delete it from the graph. After that, we continue using Shortest Path to find a new connection from A to B on new graph. This algorithm will be terminated if we cannot find any connections from A to B or we achieve the limited number of connections with each demand. Finally, we have a database with ID and corresponding connection from A to B. These connections are the solutions of a demand from A to B. The total time of CDE for each demand is $O(n^2)$.

In A-SNDP problem, there are many customers' demands, so we implement FindOtherWays routine in turn with each demand. Then, a set of databases are created, each of them holds the solutions that represent possible connection for each demand.

3.2 CDE for Solving A-SNDP

In A-SNDP, we must find two connections for each demand to guarantee the network survivability. Those are working path and backup path. In unicast demand, working path and backup path have the same source node and destination node but different destination node in the anycast demand. We will apply CDE for encoding graph in A-SNDP to build working path and backup path database.

With this building, working path and backup path do not have common edges. So if the working path breaks, the backup path will replace and restore data on the working path easily. Thus, the network survivability is always guaranteed for each demand with single failure case. Each demand in A-SNDP will have working path and backup path database, and they relate together through the foreign key. Figure 1 presents the working path and backup path databases of demand 1.

3.3 Individual Representation

In this section, we propose a new genetic algorithm using CDE for solving A-SNDP (called CDE-GA) that combine CDE with integer encoding for representing individual. An individual is represented by a couple of chromosomes with their length is equal to the number of input demands.

3.4 Fitness Function

To build the fitness function, at first, the value of bandwidth attribute is set with 0 for all links. Then, we browse over its chromosomes and add the value of required bandwidth to bandwidth attribute of chosen links. Finally, we browse over all connections to calculate the value of cost attribute of all links. Fitness function is defined as the network cost.

3.5 Genetic Operators

Crossover operator: One-point crossover is used. We choose 4 randomly individuals in current population, sort them in ascending order of fitness value, assign them in turn with X_1, X_2, X_3 and X_4 where $F(X_1) > F(X_2) > F(X_3) > F(X_4)$.
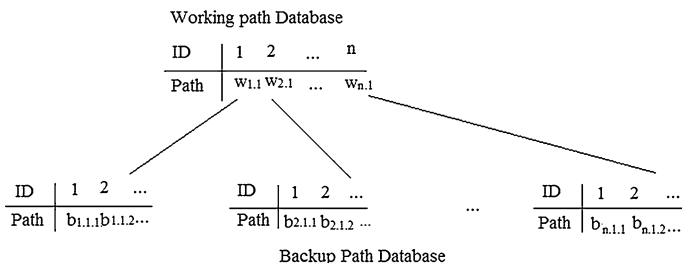
Fig. 1 The demonstration of CDE for demand 1, where $w_{i,1}$ is the working path, $b_{i,1,j}$ is the backup path of working path $w_{i,1}$, $i = 1 \to n$; j is the ordered number for backup path database

Then, X_1 and X_2 are combined into X_5, X_3 and X_4 are combined into X_6. Individuals X_5, X_6, X_3 and X_4 and replace 4 worse individuals in the population to make new population.

Mutation operator: A gene is randomly chosen in chromosome, then its value is replaced with a different value in corresponding database with this gene.

3.6 Proposed Genetic Algorithm

In this algorithm, we initialize a population randomly. Our crossover operator is one—point model which is presented by symbol $\odot$ in pseudo code. After crossover, the mutation process is implemented k times depending on the mutation rate. Gbest is best value in the population and is updated in each generation and returned after the termination condition is satisfied. After a fixed number of generations or when the result is not improved after a certain number of generations, we terminate the algorithm. The proposed algorithm is presented below.

```
1.  Procedure Genetic_Algorithm
2.  Begin
3.  pop ← INITIALPOPULATION
4.  Gbest ← best individual in the population
5.  While !Termination do begin
6.  For i = 1 to number_crossover_operation do begin
7.     Xᵢ, Yᵢ, Zᵢ, Tᵢ ← RandomSelection(pop)
           //F(Xᵢ) > F(Yᵢ) > F(Zᵢ) > F(Tᵢ)
8.        Xᵢ' ← Xᵢ ⊙ Yᵢ
9.        Yᵢ' ← Zᵢ ⊙ Tᵢ
10.       Select Xᵢ', Yᵢ' //Select for the next population
11.   Endfor
12.   For j = 1 to number_of_mutation_operations do begin
13.      Xⱼ ← RandomSelection(pop)
14.      Xⱼ' ← Mutation(Xⱼ)
15.   Endfor
16. Update GBest
17. Endwhile
18. Return GBest
```

4 Experiment Result

4.1 Experiment Setup

The proposed algorithm are evaluated and compared with Tabu Search [1], FBB1 and FBB2 [8] on three real network topologies (Polska, Germany, Atlanta

networks, see http://sndlib.zib.de/home.action). For each network instance, 10 test sets are randomly created with different content of customers' demands. In CDE-GA, the population size is 300, the number of generation is 20, giving as total maximal number of fitness's evaluation of 6,000. The mutation rate is 1 %, crossover rate is 50 %. Each test set is run 50 times. All the programs are run on a PC with Intel Core 2 Duo U7700, RAM 2 GB.

4.2 Computational Results

Figure 2 shows that on the Polska network—the smallest instance, the best results found by CDE-GA are better than FBB1 and Tabu Search, but they are worse than FBB2 in almost test sets. Figures 3 and 4 show that on the bigger instance (Germany and Atlanta), the best results found by CDE-GA are much better than the ones found by FBB1, FBB2 and Tabu Search. It means that CDE-GA can find the solutions more effectively in the big instance.

Figure 5 shows that on the Polska network, the average results of CDE-GA, Tabu Search, FBB1 and FBB2 are almost equivalent. However, as seen in Fig. 2, the average results found by CDE-GA is better than the ones found by FBB1 and Tabu Search but worse than the ones found by FBB2. Figures 6 and 7 show that on the bigger network (Germany and Atlanta) the average results found by CDE-GA are much better in comparing with Tabu Search, FBB1 and FBB2.

Figures 8 and 9 show that on small network (Polska, Germany network), the average running time of CDE-GA is longer than Tabu Search and FBB1, FBB2. Figure 10 shows that on bigger network (Atlanta network), the average running time of CDE-GA, Tabu Search, FBB1 and FBB2 are not different.
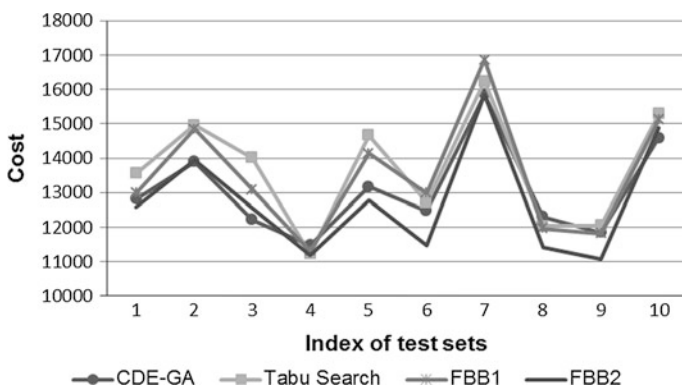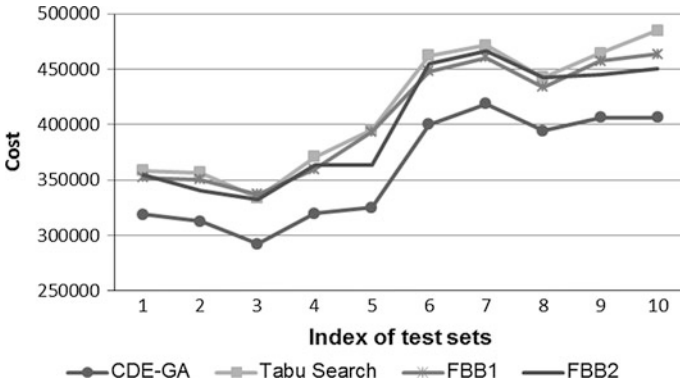

Fig. 2 The best result of Polska network
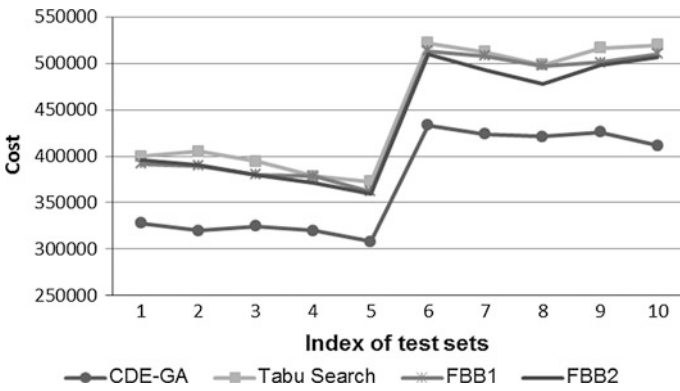
Fig. 3 The best result of Germany network

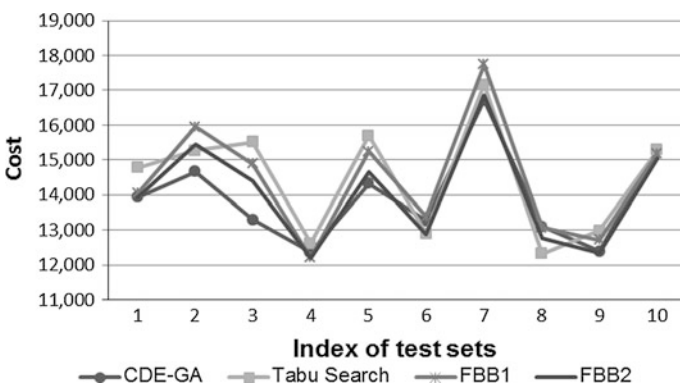**Fig. 4** The best result of Atlanta network

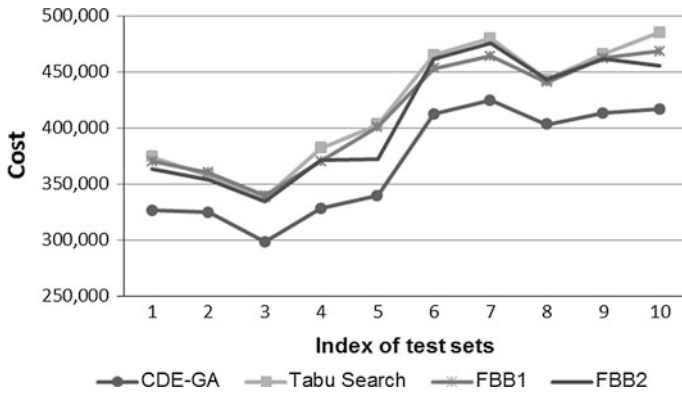**Fig. 5** The average result of Polska network

Fig. 6 The average result of Germany network
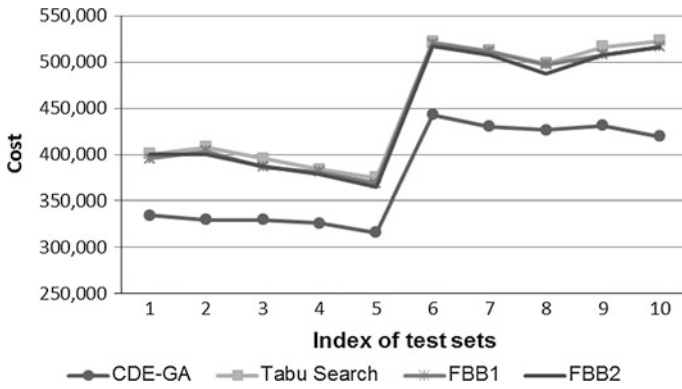

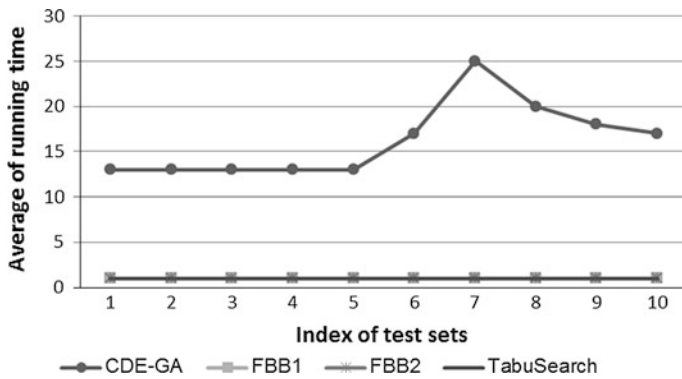**Fig. 7** The average result of Atlanta network

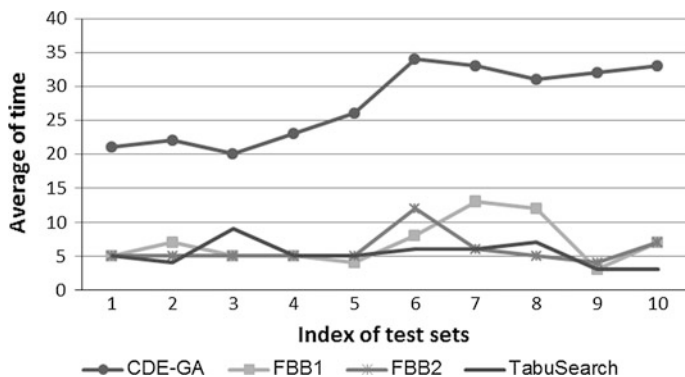**Fig. 8** Average running time on Polska network

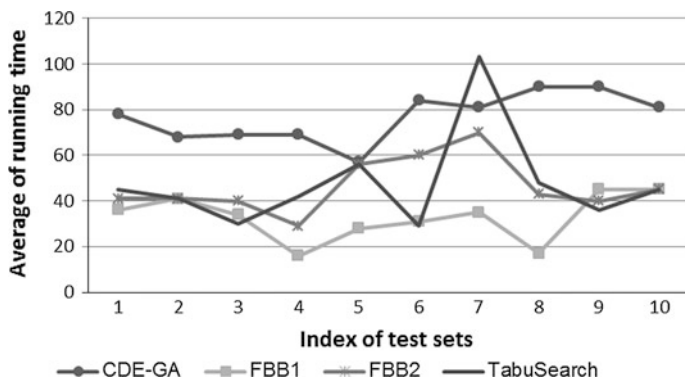**Fig. 9** Average running time on Germany network

Fig. 10 The average running time Atlanta network

5 Conclusion

In this paper, we proposed a new genetic algorithm called CDE-GA for solving A-SNDP. We experimented on three networks, for each instance, we randomly create 10 test sets which are different from the content of customers' demands. The results show that our proposed approach is quite effective for solving A-SNDP. On the big instances (Germany and Atlanta), the results found by CDE-GA are better than FBB1, FBB2 and Tabu Search. However, on the small network (Polska network), the results found by CDE-GA are only better than FBB1 and Tabu Search. In future work, we are planning to improve the algorithm for solving more bigger instances. Moreover, we hope that we can find the other approach with better results for A-SNDP.

References

1. Gładysz J et al (2010) Tabu search algorithm for survivable network design problem with simultaneous unicast and anycast flows. Int J Electron Telecommun 56(1):41-48, Versita Publisher, Warsaw
2. Walkowiak K (2008) Optimization of unicast and anycast flows in connection-oriented networks. In: Gervas O (ed) Proceedings of computational science and its applications—ICCSA 2008, LNCS, vol 5072. Springer, Perugia, pp 797–807
3. Anycast vs Unicast. http://communitydns.eu/Anycast.pdf. Accessed 18 July 2012
4. Walkowiak K (2006) A new function for optimization of working paths in survivable MPLS networks. In: Proceedings of computer and information sciences—ISCIS 2006. Springer, Istanbul, pp 424–433
5. Grover W (2004) Mesh-based survivable networks: options and strategies for optical, MPLS SONET and ATM networking. Prentice Hall PTR, New Jersey
6. Sharma V, Hellstrand F (2003) Framework for MPLS-based recovery. RFC 3469
7. Gladysz J, Walkowiak K (2009) Optimization of survivable networks with simultaneous unicast and anycast flows. In: Proceedings of ICUMT. Poland, pp 1–6
8. Binh H et al (2012) Heuristic algorithms for solving survivable network design problem with simultaneous unicast and anycast flows. In: Proceedings of 8th international conference on intelligence on computing, ICIC 2012. Huangshang, China, pp 137–145
9. Nissen V, Gold S (2008) Survivable network design with an evolution strategy. In: Proceedings of success in evolutionary computation. Springer, Berlin, pp 263–283
10. Pioro M, Medhi D (2004) Routing, flow, and capacity design in communication and computer networks. Morgan Kaufmann Publishers, San Francisco
11. Battiti R, Brunato M, Mascia F (2008) Reactive search intelligent optimization. Springer, New York
12. Vasseur J, Pickavet M, Demeester P (2004) Network recovery: protection and restoration of optical, SONET-SDH IP and MPLS. Morgan Kaufmann, San Francisco
13. Kasprzak A (1989) Algorithms of flow, capacity and topology structure in computer networks. Monography, Wroclaw
14. Walkowiak K (2003) Anycast communication—a new approach to survivability of connection-oriented networks. In: Proceedings of communications in computer and information science. Springer, Berlin, pp 378–389
15. Johnson D, Deering S (1999) Reserved IPv6 subnet anycast addresses. RFC 2526
16. Michalewicz Z (1995) Genetic algorithms + data structures = evolution programs, 3rd edn. Springer

Based on the M-GIS System Structure of the Intelligent Mobile Phone Positioning and Tracking System

YueChun Feng and HaiRong Wang

Abstract Based on the Android platform of mobile phone location tracking system will combined phone safety protection with positioning technologies, and use Google map seamlessly combine this advantage, to provide users with precise positioning service, changing the card lock machine to pre-set the phone number to send a warning message, remote control, data backup, the positioning of the machine, Friends positioning and electronic fencing. This system can be used for mobile phone anti-theft tracking, monitoring and care of the parents of children with specific financial and personnel monitoring, to provide users with a convenient and reliable service, with good commercial value and social value.

1 Introduction

At present, there are some based on the Android [1] platform of mobile phone protection methods and products, such as the Technical Research Centre of Finland (VTT), researchers have invented a way to prevent the theft of mobile phones and laptop technology, which uses biometric technology to identify the

Y. Feng (✉) · H. Wang
Department of Electrical and Information Engineering, Ningxia Institute of Science and Technology, Ningxia, People's Republic of China
e-mail: fengyuechun@163.com

H. Wang
e-mail: wanghairong@163.com

user's pace, which to determine whether the user is a legitimate user. There is also a mobile phone protection technology—"Mobile self-destruct", the principle by hardware control, making the stolen phone to lose value, but this way too much loss. Throughout the mobile phone market, mobile phone protection has two types of software and hardware, and from the hardware way or the factory-installed protective function, after all, is limited, and can not meet the needs of the majority of mobile phone users for mobile phone protection. The introduction and development of the smartphone protection software provides a platform can be achieved. Smart phone provides a standard platform, so any third-party software used in intelligent systems can be developed.

2 Principle and Architecture Diagram and Function of the System

2.1 System Principle

The smart phone by receiving GPS [2] data is calculated to determine the phone's location information (longitude, latitude, time, trajectory parameters), and report the results to the mobile network. Accuracy with GPS, up to 5–10 m. Receiving satellite positioning information and calculate the current location information, internal communication module using the TCP/IP protocol such information is automatically transmitted to the designated public IP address monitoring and dispatching center communication server (IP address, upload interval parameters in advance phone settings), communication server receives phone location information and save it into the database, the PC access to the database server, you can phone query.

2.2 System Functions

System Settings

System registration function for the initial use of the phone in the entire system, the user simply fill in the user name and password for two, after the submission of the registration command, the system extracts the phone's IMEI number, SIM card number, phone number, this information is sent to the server, create a new user profile. GPS settings, GPS start and off by the user to complete the upload interval can be set to 5, 10 min, can not set, do not set the default for not taking the

initiative upload. Backup settings, you can set the reservation phone number and E-mail to set aside the cell phone number used to send SMS alarm contains a replacement SIM card information lost phone, reserved mailbox for important data backup phone.

phone protection function

Security settings designed for handset protection, enter the password for the phone lock and unlock, enter custom phone number. Set a password for the first time, directly enter a password to the first set after the success of the password, first enter the old password in order to change the password, and then enter the new password twice, you can change the password successfully submitted. Enter a custom phone numbers to the phone is lost, others replace the user's mobile phone SIM card, you can send a short message to the number of custom.

positioning tracking function

Phone client via GPRS channel information (latitude and longitude, phone model, screen size, etc.) of the machine is sent to the server, the server will get the local coordinates of the location coordinates and user are reflected in the map. Users can also customize the trigger prompts for the location of the Friends. For example, parents can customize the location of points of interest in their children, parents worry about their children to cafes and other places, cafes, game room location can be set to the point of interest, when their children's mobile phone location report in cafes near to parents sent phone messages to tell their parents of their friends have been in the range of the points of interest. Parents also can set a point of interest for the school, and will automatically prompt information to parents when their children out of school.

3 Subject Research Method Process

3.1 M-GIS System Architecture

M-GIS [3] include three levels of the presentation layer, logic layer, data layer, as shown M-GIS system based on the Android platform architecture, the presentation layer, including Android phones, handheld computers, vehicle terminal is equipped with the Android platform mobile devices, logic layer contains the wireless gateway, Web server and GIS server, data layers including spatial database with the local database. As shown.

3.2 Technology Roadmap

Changing the card lock machine

The phone is switched on after detecting the SIM card information is replaced, if the replacement will add the boot the matters lock screen, the user has only to enter the correct password for successfully unlock continue to use the phone.

Remote Control

Phone is lost, the user can preset mobile phone number on the phone to the remote control system based on user-defined protocol detects the user issued password eligibility, if qualified accordingly.

Step1: phone is lost (which may be placed temporarily can not remember the place, it could be stolen);
Step2: According to the protocol to use, access to the specific location of the phone through the binding number of friends to send text messages;
Step3: according to the use agreement through binding number of friends to get phone contacts, SMS, call history, photos, audio, video, etc.;
Step4: according to the use agreement, by binding Friends No. delete phone contacts, SMS, call history, photos, audio, video.

Data Backup

Lost phone to the rightful owners for privacy and contacts in the phone is the most important and the most dangerous of all, it is very necessary to delete the information in the phone so in the limited time offsite backup phone information. Steps as follows:

Step1: obtain initialization settings of a contact name, telephone and E-mail
Step2: Upload data

Locate

The GPS technology LBS (Location Based Services, location-based services) research. LBS has two meanings: The first is to determine the geographic location of the mobile device or user; Second is to provide all kinds of information services and location. Means a wide range of services and positioning system, referred to as "location-based services, so that the LBS is to make use of the Internet or wireless network between a fixed or mobile users, the positioning and services. GPS positioning applications of this study, is also based on the idea of the LBS

positioning tracking function. Use before Friends positioning function, users and friends of users of the machine must be a registered user of the system, the information submitted by the phone's the uniquely identifies IMEI IMEI, SIM card number, phone number, this information by phone first to submit registration, and then through registered to fill in the rest of the information, such as name, address, other phone, friends name, friends phone number, and so on. Functions are as follows:

1. GPS function is activated, the first to calculate the latitude and longitude of the phone's location.
2. Only the user manual operation, the positioning information back the Phone UI displayed to the user.
3. Sent to the server, the location information is packaged and sent to the server via GPRS.
4. If the user chooses to locate friends, buddy's phone number, and Friends of the requests made to the server queries. The server determines inquiries with the person making the inquiry is a friend relationship, if confirmed friends relationship server to query to send GPS positioning command; returned to the person making the inquiry, if not friends relationship, a query ultra vires information can not query.
5. Query the person receives the server's command to start the GPS function, locate their own position, then returns the longitude and latitude information via GPRS to the server, and then returned by the server to be queried.
6. Friends user receives positioning information packet, calculate the distance and orientation between the two, and prompts the user to have been found or positioning information, the user to view the result of the calculation.
7. Server returns the location map, to mark both positioning information can be viewed on a map.

Electronic fence

Trigger prompts the essence of the technology is the position of the electronic fence, a virtual area can be set by software or hardware in when monitoring objects enter or leave a particular area can alarm to the server, the server administrator or responsible person can be notified according to the actual needs. This technology can be applied to monitor and custody of the parents of children, parents customized information of the location of points of interest (such as an Internet cafe), this information back to the children on the phone, in the form of files stored when the child user to reach this position, GPS [4] positioning information back to the server, the information sent by the server to notify the parents.

4 Results and Conclusions

4.1 Achieving Results

Acquisition of basic phone information

The realization of the user registration function is divided into two parts, first, to obtain the SIM card number, IMEI number, manufacturer, phone model information; Second, to obtain the information package, together with the phone number is sent to the remote server, registration.

Changing the card lock machine

When the system detects a mobile phone SIM card is changed as soon as the lock screen, the user will need to enter the correct password to unlock.

Data Backup

The Android platform address book data stored in the database of the data table contacts.db, which saved a contact name, phone number, e-mail, notes and other information. Backup concrete realization of these data as follows:

Step1: get through the the ContentResolver object's query (Phone. CONTENT_URI, PHONES_PROJECT, null, null, null) method phone associates Cursor object phone Cursor.

Step2: phoneCursor.getString(PHONES_NUMBER_INDEX), contact name, mailbox access is also true, just the parameters are different.

Step3: get data connection into a string and converted into binary data, by means of the HttpURLConnection set context type and length, and then output.

Alarm sound control

1. play alarm sound

When binding friends number send command code is "008", have to perform this function, to realize the function of the method used for: (1) MediaPlayer mp = MediaPlayer.create(context, R.raw.alarm2); (2) mp.setLoping(true); (3) mp.start();

2. shut off the alarm sound

When binding friends number send command code is "009", have to perform this function, to realize the function of the method used for:
Alarm sound control (1) mp.setLooping(false); (2) mp.stop(); (3) mp.release();

Remote tracking

Mobile phone is lost real-time positioning its own position, and the position of the latitude and longitude information sent to the preset phone number, then the user

can according to the latitude and longitude information in web server find out the location of the mobile phone is lost.

1. to the preset phone number send position information:

```
SmsManager smsManager = SmsManager.getDefault();
PendingIntent mPI = PendingIntent.getBroadcast(context,0,new Intent(), 0);
String strDestAddress = "";
String strMessage = "经度:"+location.getLatitude()+"\n纬度"+location.get-
Longitude();
    phoneDbAdapter = phoneDbAdapter.open();
    if (phoneDbAdapter != null) {strDestAddress= phoneDbAdapter.fetch-
PhoneKey(PhoneDbAdapter.KEY_BACK_NUMBER);
  }
    phoneDbAdapter.close();
    if (strDestAddress != null && !strDestAddress.equals("")) {
        ArrayList<String>texts=smsManager.divideMessage(strMessage);
    for(Stringtext  :texts){smsManager.sendTextMessage(strDestAddress,  null,
text, mPI, null);
      }
}
```

2. through the accurate longitude and latitude information in Google maps inquires the mobile phone is lost the current position.

4.2 Verification Results

1. detection is legitimate user login, and in the legal user login to corresponding security Settings, such as open GPS positioning function, set up the data backup email address, and phone number, etc.
2. in the GPS positioning service open cases can realize positioning, and grab the current position of the latitude and longitude.
3. can be sent through accord with user-defined protocol SMS remote control mobile phone is lost, such as sending HH101CMD008123 can make the mobile phone is lost play alarm bell.

5 Problems and Discussed

1. location function
2. the automatic alarm function
3. the automatic alarm module and public service network connected

References

1. Gadri RC, Alhat B, Chavan A, Kamble S, Sonawane R (2012) Land vehicle tracking system using Java on android platform. In the Computer engineering and intelligent system
2. Lu M, Chen W, Shen X, Lam HC, Liu J (2007) Positioning and tracking construction vehicles in highly dense urban areas and building construction sites. In the Automation in Construction
3. Tsou MH (2004) Integrated mobile GIS and wireless internet map servers for environmental monitoring and management. Cartography Geogr Inf Sci
4. Goh KN, Ng YP, Jusoff K, Chen YY, Tan YY (2011) Architecture of a GPS-based road management system. World Appl Sci J 12(Special Issue on Computer Applications & Knowledge Management)

A Comparison of Actual and Artifactual Features Based on Fractal Analyses: Resting-State MEG Data

Montri Phothisonothai, Hiroyuki Tsubomi, Aki Kondo, Yuko Yoshimura, Mitsuru Kikuchi, Yoshio Minabe and Katsumi Watanabe

Abstract Future standardized system for distinguishing actual and artifactual magnetoencephalogram (MEG) data is an essential tool. In this paper, we proposed the quantitative parameters based on fractal dimension (FD) analyses in which the FD may convey different features before and after artifact removal. The six FD algorithms based on time-series computation, namely, box-counting method (BCM), variance fractal dimension (VFD), Higuchi's method (HM), Kazt's method (KM), detrended fluctuation analysis (DFA), and modified zero-crossing rate (MZCR) were compared. These approaches measure nonlinear-behavioral responses in the resting-state MEG data. Experimental results showed that the FD value of actual MEG was increased statistically in comparison with the artifactual MEG. The DFA and the HM present a best performance for analyzing simulated data and resting-state MEG data, respectively.

Keywords Fractal dimension · Complexity measure · MEG · Nonlinear analysis · Artifact removal

M. Phothisonothai (✉) · A. Kondo · K. Watanabe
Research Center for Advanced Science and Technology, The University of Tokyo,
Tokyo 153-8904, Japan
e-mail: montri@fennel.rcast.u-tokyo.ac.jp

A. Kondo
e-mail: kondo@fennel.rcast.u-tokyo.ac.jp

K. Watanabe
e-mail: kw@fennel.rcast.u-tokyo.ac.jp

H. Tsubomi
Faculty of Humanities, The University of Toyama, Toyama 930-8555, Japan
e-mail: htsubomi@hmt.u-toyama.ac.jp

Y. Yoshimura · M. Kikuchi · Y. Minabe
Research Center for Child Mental Development, Graduate School of Medical Science
Kanazawa University, Kanazawa 920-8641, Japan
e-mail: minabe@med.m.kanazawa-u.ac.jp

1 Introduction

Magnetoencephalogramphy (MEG) is a non-invasive technique for investigating human brain activities through magnetic field measurement. However, one fundamental problem is that the presence of physiological artifacts could interfere with the measurement during recording. Recently, automatic methods for artifact removal have been studied based on independent component analysis (ICA). These effective methods can eliminate the artifactual components from raw MEG recording [1–4]. However, the classification of actual and artifactual MEG makes analysis much more difficult due to many unknown variables. Thus, this exploratory study is to propose quantitative parameter for identifying what is *an actual MEG* and what is *an artifactual MEG* data in the human brain in terms of nonlinear-behavioral study which may convey useful information and showing a strong correlation with the physiological data [5]. Measures of irregularity and complexity in time-series data can be quantified by using fractal dimension (FD). Waveform of FD values indicates the complexity of a pattern or the quantity of information embodied in a waveform pattern in terms of morphology, spectra, and variance. The applicability of fractal analysis on physiological time-series data has been examined in various studies [6], e.g., electroencephalogram (EEG) epilepsy, lung-sound analysis, MEG Alzheimer's disease, near-infrared spectroscopy (NIRS), EEG drowsiness, and speech. Unfortunately, although precise methods to determine the FD have already been proposed, their usefulness is severely limited since they are computer intensive and their evaluation especially for the short-time series analysis [7].

The main objective of the present work, in this paper, we compared the different performance of FD algorithms using simulated fractional Brownian motion (fBm) and synthetic (Syn) data in which algorithms are implemented in time domain based computation. Especially, this study focuses on the length of 2^{10}-point, moving-overlapped interval of 2^8-point, and evaluates six FD algorithms in terms of their ability and variability in estimating FD value. The best performance of FD algorithm and the quantitative parameters for distinguishing actual and artifactual MEG data are discussed.

2 Fractal Dimension Analysis

A method of estimating dimension can be determined on the basis of time and frequency domains. In this paper, we focused on time domain based FD algorithms since they are more effective than frequency domain based algorithms [8]. The six FD algorithms can be defined as follows:

2.1 Box-Counting Method

Classical one of the most common methods for calculating the FD of a self-similar fractal is the box-counting method (BCM). The definition of BCM is a bounded set in the Euclidian n–space that it composes self-similar property [9], $N_r \propto r^{-D}$ by covering a structure with boxes of radii, r, the FD can be determined by

$$D_{BCM} = \lim_{r \to 0} \frac{\log_2(N_r)}{\log_2(1/r)}, \tag{1}$$

this process with several different radii is repeated. To implement this algorithm, number of contained box, n_r, is computed from the difference between the maximum and minimum amplitudes of the data divide by the changed radius $n_r(i) = \lfloor (\max(x_r) - \min(x_r))/r(i) \rfloor$ and a total number of contained boxes $N_r = \sum n_r(i)$ where x_r represents the input time series with length L, $r(i)$ is a radius by changing a step of k within the i-th subdivision window, and integer-part function denoted by $\lfloor \cdot \rfloor$. To determine the FD, least-square linear fitted line corresponds to the slope of the plot $\log_2(N_r)$ versus $\log_2(1/r)$ is applied. In this study, we set $r_{max} = 2^9$.

2.2 Variance Fractal Dimension

The basic idea of calculation is based on the power law relationship between the variance of the amplitude increments of the input time series, which was produced by a dynamic process over time. The main advantage of variance fractal dimension (VFD) was its support of the real-time computation [10]. The amplitude increments of a datum over a time interval Δt adhere to the following power law relationship $\mathrm{var}\,[x(t_2 - t_1)] \propto |t_2 - t_1|^{2H}$, the Hurst exponent can be calculated by using a log–log plot then given by

$$H = \lim_{\Delta t \to 0} \left(\frac{1}{2} \frac{\log_2[(\Delta x)_{\Delta t}]}{\log_2(\Delta t)} \right), \tag{2}$$

the variance in each window per stage k can be calculated as follows

$$\mathrm{var}\,[\Delta x_{\Delta t}] = \frac{1}{N_k - 1} \left[\sum_{j=1}^{N_k} (\Delta x)_{jk}^2 - \frac{1}{N_k} \left(\sum_{j=1}^{N_k} (\Delta x)_{jk} \right)^2 \right], \tag{3}$$

where $(\Delta x)_{jk} = x(jn_k) - x((j-1)n_k)$. The least-square linear fitted line corresponds to the slope of the plot $\log_2(n_k)$ and $\log_2(Var[\Delta x]_k)$, the Hurst exponent is computed as $H = 0.5s$ where s is the obtained slope then the FD can be estimated as $D_{VFD} = 2 - H$. The values k represents the integer range chosen such that each window of size N_T contains a number $n_k = 2^k$ of smaller windows of size $N_k = \lfloor N_T/n_k \rfloor$. In this study, we set $k_{max} = 8$.

2.3 Higuchi's Method

Higuchi proposed an algorithm for the estimation of FD directly in time domain without reconstructing the strange attractor [11]. The Higuchi's Method (HM) algorithm based on the given finite time series $y = \{y(1), y(2), \ldots, y(N)\}$, a new time series, y_m^j, are constructed by the following equation $y_m^j = \left\{y(m), \ldots, y\left(m + \left\lfloor \frac{N-m}{j} \right\rfloor\right) j\right\}$ where both m and j are integers and they indicate the initial time and the time interval, respectively. The length, $L_m(j)$, of each curves is computed as

$$L_m(j) = \frac{1}{j}\left\{\left(\sum_{i=1}^{\left\lfloor \frac{N-m}{j} \right\rfloor} \left| \begin{array}{l} y(m+ij)- \\ y(m+(i-1)j) \end{array} \right| \right) p \right\} where\, p = \frac{N-1}{\left\lfloor \frac{N-m}{j} \right\rfloor j}. \quad (4)$$

The length of curve for time interval j, $L_m(j)$ is computed as the average of the m curves. A relationship of this algorithm is $L_m(j) \propto j^{-D_{HM}}$ therefore we apply the least-squares fitting line of $\log_2 L(j)$ versus $\log_2 j$, the negative slope of the obtained line is calculated giving the estimate of the FD, D_{HM}. In this study, we set $j_{max} = 2^6$.

2.4 Katz's Method

Katz's FD estimation is derived directly from the waveform, eliminating the preprocessing step of creating a binary sequence [12]. The FD of a curve can be defined as $D_{KM} = \log L / \log d$ where L is the total length of the curve or sum of distances between successive points, and d is the diameter estimated as the distance between the first point of the sequence and the point of the sequence that provides the farthest distance. Mathematically, d can be expressed as $d = \max_{i \in L}\left\{\sqrt{1 + (x(i) - x(i-1))^2}\right\}$, considering the distance between each point of the sequence and the first, point i is the one that maximizes the distance with respect to the first point. Evaluating FD by Katz's method (KM), the average step or average distance between successive points, $\bar{a}$. Normalizing distances by defining the number of steps in the curve, $n = L/\bar{a}$, then can be written as $D_{KM} = \log_2(n)/(\log_2(\frac{d}{L}) + \log_2(n))$.

2.5 Detrended Fluctuation Analysis

This method was initially proposed by Peng et al. [13]. Therefore, the FD estimation using detrended fluctuation analysis (DFA) can be defined as $X(k) = \sum_{i=1}^{k} [x(i) - \bar{x}]$ this integrated series is divided into non-overlapping intervals of length n. In each interval, a least squares line is fit to the data. The series $X(k)$ is then locally detrended by subtracting the theoretical values $X_n(k)$ given by the regression. For a given interval length n, the characteristic size of fluctuation for this integrated and detrended series is calculated by $F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^{N} [X(k) - X_n(k)]^2}$. This computation is repeated over all possible interval lengths (in practice, the minimum length is around 10, and the maximum is a half length of input data, giving two adjacent intervals). A relationship is expected, as $F(n) \propto n^{\alpha}$ where α is expressed as the slope of a double logarithmic plot $\log_2(F(n))$ versus $\log_2(n)$. Then α can be converted into the Hurst exponent $H = \alpha - 1$ and the estimated FD according as $D_{DFA} = \alpha - 3$. In this study, we ranged $k = 4, 5, ..., 9$.

2.6 Modified Zero-Crossing Rate

Phothisonothai and Nakagawa proposed the FD estimation method on the basis of zero-crossing function or called modified zero-crossing rate (MZCR) [14]. The proposed MZCR can be defined on the general relationship of power law, during the locally computation period of input data as $f_z(\Delta x) \propto \Delta t^{\mu}$, where $f_z(\cdot)$ is the MZCR function and μ is a scaling parameter. The input time-series data are formed in the length of 2^m-point. To determine FD value by means of MZCR method, there are three main steps of processing as follows: step (1) zero-mean data by subtracting the mean value of locally sampled period from each value as $\bar{x}$, step (2) the zero-mean data is then locally detrended by the linear regression, y_r, as $\hat{y} = \bar{x} - y_r$, and step (3) the ZCR value for those detrended data could be determined by $f_z(m) = \frac{1}{L} \sum_{m=0}^{L-1} \Phi(\hat{y}_m \hat{y}_{m-1})$ where $\Phi(\hat{y}_m \hat{y}_{m-1}) = 1$ if $\hat{y}_m \hat{y}_{m-1} < 1$ otherwise $\Phi(\hat{y}_m \hat{y}_{m-1}) = 0$. The estimated FD is then determined by $D_{MZCR} = 2 + \mu$. This computation is repeated over all possible interval lengths. We can determine the scaling parameter, μ, by the least-squares fitting line of $\log_2 f_z(m)$ versus $\log_2 m$, the slope of the obtained line is calculated giving the estimate of the FD. In this study, we set $m = 2^2, 2^3, , 2^9$.

3 Experiment and Results

3.1 Resting-State MEG Data Acquisition

MEG data were recorded with a 151-channel superconducting quantum interference device (SQUID) whole-head coaxial gradiometer MEG system (PQ 1151R; Yokogawa/KIT) in magnetically shielded room (Diode Steel) [15]. MEG data were digitized by the sampling rate of 1 kHz 16-bit resolution. Ten healthy children (5 boys, 5 girls)[1] participated in this experiment. The participants had a mean chronological age of 65.4 months. All subjects were given to perform the resting state for three minutes. During MEG recording, children lay on the bed and viewed a cartoon animation programs with stories through a TV monitor especially attractive to young children. Off-line analysis of MEG data was performed with Yokogawa MEG Reader Toolbox Rev1.4 and Matlab (MathWorks).

3.2 Resting-State MEG Data Analysis

To present the different performances of each FD algorithm, we firstly assessed FD algorithms with artificially generated two data that FD value can be easily set as follows: (1) fBm data. In this study, the fBm data were simulated by a discrete zero-mean stationary Gaussian process which proposed by Kroese and Botev [16], and (2) Syn data. Syn data were simulated by discrete Weierstrauss cosine function which can be defined as $W_h(n) = \sum_{i=0}^{M} \gamma^{-iH} \cos(2\pi\gamma^i n)$ where $\gamma > 1$, we set $M = 26$ and $\gamma = 5$. The FD is given by $D = 2 - H$. This experiment, we tested the fBm and Syn data 2^{10}-point length by setting step-size FD resolution of 0.001 and iteration of 50 runs. The output performances are shown in Fig. 1. To show the accuracy and variability, we compared the estimated FD with the theoretical FD value (true FD value) by using two basic statistical parameters, i.e., mean-squared error (MSE) value and standard deviation (SD).

For the analysis of resting-state MEG data, the raw MEG data were assumed as artifactual data and the cleaned MEG data were assumed as actual data throughout the experiment. Such artifacts in raw MEG data were firstly removed by means of our method [4]. Two minutes of a whole resting state, i.e., 30 s to 1 min 30 s, has been examined. Then, we employed the time-dependent fractal dimension (TDFD) approach [17] in which the trajectory of FDs can easily be revealed. The number of possible points is defined as $N_d = \lfloor (L - L_w)/\Delta t \rfloor + 1$ where L is a total length (120,000 samples), L_w is a window length (or frame) $L_w \leq L$, and Δt is an interval,

---

[1] All subjects had normal brain function and data collection was approved by the Ethics Committee of Kanazawa University Hospital, all which were in accordance with the Declaration of Helsinki.
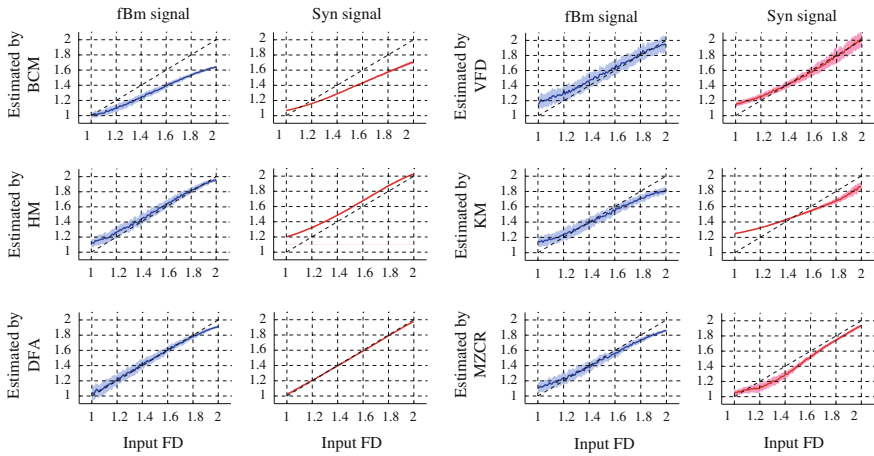
Fig. 1 Performance comparison of the input theoretical FDs versus the estimated FDs for fBm and Syn data
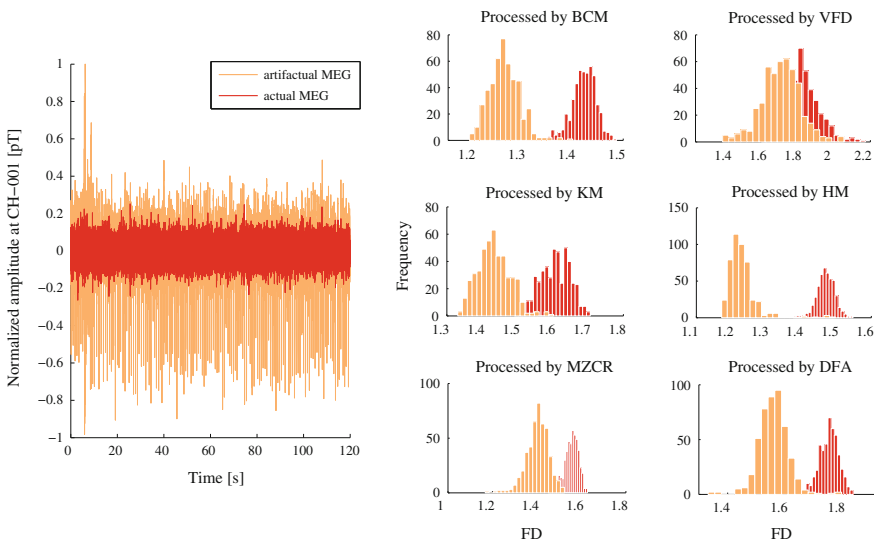
Fig. 2 (*Left*) Typical artifactual MEG plotted in *orange line* and actual MEG plotted in *red line*; (*Right*) histogram plots of all FD distributions using TDFD where a fixed-bin number is 20

we set a window function is a rectangular type, $L_w = 2^{10}$ with moving-overlapped interval $\Delta t = 2^8$. The total length $N_d = 465$ points of FD with respect to time. The analyzed results by using TDFD are shown in Fig. 2. As the correlation feature of TDFD changes between before and after artifact removal, the histogram based mapping was selected. By this approach, the frequency of occurrences of N_d-point can be computed as probability density value where measures the mutual

Table 1 Comparison results of all quantitative parameters using FD as feature

Parameter	BCM	VFD	HM	KM	DFA	MZCR
dEV	+0.1586	+0.1151	+0.2389	+0.1723	+0.1928	+0.1531
SD of TDFD (artifactual)	0.0300	0.1112	0.0420	0.0455	0.0515	0.0504
SD of TDFD (actual)	0.0240	0.0986	0.0256	0.0434	0.0345	0.0306
p–value	<0.001	<0.01	<0.001	<0.001	<0.001	<0.001
Joint probability	0.0151	0.6496	0.0196	0.0410	0.0302	0.0412
MSE (fBm)[a]	4.2101	0.5846	0.2453	0.6842	0.0803	0.3813
Sum of SD-squared (fBm)[a]	0.1381	0.9646	0.3527	0.4543	0.3581	0.3428
MSE (Syn)[a]	2.7110	0.2626	1.0289	1.4111	0.0123	0.5864
Sum of SD-squared (Syn)[a]	0.2626	0.4241	0.0141	0.0671	0.0134	0.1493
Average computation time[b]	≈5.2	≈0.8	≈11.8	≈0.7	≈16.1	≈91.4

[a]Results obtained from Fig. 1

[b]Performed with processor Intel Core Duo 1.5 GHz, memory 8 GB, unit in ms.

dependence of the two random variables. Accordingly, the different expected value (dEV) can be evaluated by subtracting the EV of actual data from the EV of artifactual data. Table 1 shows the comparison results obtained by this experiment which these results were calculated as grand value from all subjects and all channels.

4 Discussions and Conclusions

Our experiment was to identify feature of resting neuromagnetic activity from MEG data by means of FD analyses. Based on the obtained results in Table 1, the research discussed in this contribution contains two main aspects: on the best performance of FD algorithm and on the quantitative parameters for distinguishing actual and artifactual MEG data. We found that the DFA algorithm presents a best performance, i.e., lowest MSE of $0.0803(\pm0.3581)$ and lowest MSE of $0.0123(\pm0.0134)$ for fBm and Syn data, respectively. In computing, the KM and the VFD are faster than others in terms of time consumption, i.e., less than a millisecond. We also found that the HM presents a best performance in terms of maximum distance dEV = 0.2389 with reasonable joint probability of 0.0196 as shown in Fig. 2. All FD algorithms except VFD showed that the difference to be statistically significant $p < 0.001$ by using one-way analysis of variance (ANOVA). The main contribution of this study was to present optimal FD algorithm and the quantitative parameters for distinguishing between actual and artifactual MEG data. The results may be especially useful for analyzing and identifying complex biological waveforms, i.e., resting-state MEG data.

References

1. Mantini D, Franciotti R, Romani GL, Pizzella V (2008) Improving MEG source localizations: an automated method for complete artifact removal based on independent component analysis. NeuroImage 40:160–173
2. Mantini D, Penna SD, Marzetti L, de Pasquale F, Pizzella V, Corbetta M, Romani GL (2011) A signal-processing pipeline for magnetoencephalography resting-state networks. Brain Connectivity 1(1):49–59
3. Rong F, Contreras-Vidal JL (2006) Magnetoencephalographic artifact identification and automatic removal based on independent component analysis and categorization Approaches. J Neurosci Methods 157:337–354
4. Phothisonothai M, Tsubomi H, Kondo A, Kikuchi M, Yoshimura Y, Minabe Y, Watanabe K (2012) Linear and nonlinear features for automatic artifacts removal from MEG data based on ICA. In: 4th APSIPA ASC, IEEE press (to appear), Los Angeles, USA
5. Hornero R, Escudero J, Fernández A, Poza J, Gómez C (2008) Spectral and nonlinear analysis of MEG background activity in patients with Alzheimer's disease. IEEE Trans Biomed Eng 55(6):1658–1665
6. Stam CJ (2005) Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. Clin Neurophysiol 116:2266–2301
7. Delignieres D, Ramdani S, Lemoine L, Torre K, Fortes M, Ninot G (2006) Fractal analyses for 'Short' time series: a re-assessment of classical methods. J Math Psychol 50:525–544
8. Phothisonothai M, Nakagawa M (2007) Fractal-based EEG data analysis of body parts movement imagery tasks. J Physiol Sci 57(4):217–226
9. Mandelbrot BB (1983) The fractal geometry of nature. WH Freeman, New York
10. Kinsner W (1994) Batch and real-time computation of a fractal dimension based on variance of a time series. Technical report, del946. University of Manitoba, Canada
11. Higuchi T (1988) Approach to an irregular time series on the basis of the fractal theory. Physica D 31:277–283
12. Katz MJ (1988) Fractals and the analysis of waveforms. Comput Biol Med 18(3):145–156
13. Peng CK, Buldyrev SV, Havlin S, Simons M, Stanley HE, Goldberger AL (1994) Mosaic organization of DNA nucleotides. Phys Rev E 49:1685–1689
14. Phothisonothai M, Nakagawa M (2008) A complexity measure based on modified zero-crossing rate function for biomedical signal processing. In: 13th international conference on biomedical engineering (ICBME2008), vol 23. Springer Press, Singapore, pp. 240–244
15. Kikuchi M et al (2011) Lateralized theta wave connectivity and language performance in 2- and 5-Year-Old children. J Neurosci 31(42):14984–14988
16. Kroese DP, Botev ZI (2013) Spatial process generation: lectures on stochastic geometry, spatial statistics and random fields. In: Schmidt V (ed) Analysis, modeling and simulation of complex structures, vol II. Springer (to appear), Berlin
17. Sabanal S, Nakagawa M (1995) A study of time-dependent fractal dimension of vocal sound. J Phys Soc Jpn 64:3226–3238

Erratum to: A New Parallel Genetic Algorithm Based on TriBA Topological Structure

Kang Sun and Chang Wook Ahn

Erratum to:
Chapter "A New Parallel Genetic Algorithm Based on TriBA Topological Structure" in: Z. Yin et al. (eds.),
Proceedings of The Eighth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA), 2013, **DOI 10.1007/978-3-642-37502-6_61**

The second author name of the chapter "A New Parallel Genetic Algorithm Based on TriBA Topological Structure" Wook Ahn Chang should read as Chang Wook Ahn.

K. Sun (✉) · C. W. Ahn
Department of Computer Engineering, Sungkyunkwan University (SKKU),
2066 Seobu-ro, Suwon 440-746, Republic of Korea
e-mail: sunkangskku@gmail.com

C. W. Ahn
e-mail: cwan@skku.edu

Retraction Note to: Using Inner-Outer Factorization to Solve the Spectral Factor of Discrete-Time Descriptor Systems

Luhua Liang, Wei Xing and Rongwu Xiang

Retraction Note to:
Chapter "Using Inner-Outer Factorization to Solve
the Spectral Factor of Discrete-Time Descriptor Systems"
in: Z. Yin et al. (eds.), *Proceedings of The Eighth*
International Conference on Bio-Inspired Computing:
Theories and Applications (BIC-TA), 2013, **Advances**
in Intelligent Systems and Computing 212,
DOI: 10.1007/978-3-642-37502-6_33

The chapter published in the book 'Proceedings of The Eighth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA), 2013', pages 267–275, DOI: 10.1007/978-3-642-37502-6_33 has been retracted because it contains significant parts plagiarizing another publication: 'Using Inner-Outer Factorization to Solve the Spectral Factor of Discrete-Time Descriptor Systems', ((2005) Automatica, 41 (11), pp. 1855–1866; doi: 10.1016/j.automatica.2005.04.009).

L. Liang · R. Xiang (✉)
Staff Room of Mathematics, Shenyang Pharmaceutical University, Shenyang 110016, China
e-mail: xrwlove@163.com

L. Liang
e-mail: liangluhua1988@163.com

W. Xing
Institute of Systems Science, Northeastern University, Shenyang 11004, China
e-mail: awxing@21cn.com