2010 3rd International Conference on Computer and Electrical Engineering (ICCEE 2010)
IPCSIT vol. 53 (2012) © (2012) IACSIT Press, Singapore
DOI: 10.7763/IPCSIT.2012.V53.No.1.39

Ramulus Cinnamomi Intelligence Quality Assessment Based on Knowledge Reduction

Jianyang Lin[+], Mingyan Jiang, Yajie Xu and Zhoumi Kan

Department of Pharmacy, The First Hospital of China Medical University, Shenyang city, China.

Abstract. According to the traditional morphological classification divide the quality of traditional Chinese medicine Ramulus Cinnamomi into first grade and second grade. Disperse the chromatogram data Ramulus Cinnamomi obtained by criterion inspect and make knowledge reduction. Then it will get the largest liner irrespective vector group. It will get every clustering centre data of Ramulus Cinnamomi with adaptive weight PSO and POWELL optimization. Make the criterion data of Ramulus Cinnamomi as Weibull distribution and make the sample as Normal distribution to calculate similar and ensure the finial similar with similarity ratio. It used lobato integral and Hermite interpolation. This investigation makes good evaluation to the Ramulus Cinnamomi in Guangdong Guangxi Yunnan and Vietnam and high delicacy and anti-jamming.

Keywords: Quality Assessment, Distributed Clustering, Ramulus Cinnamomi, Knowledge Reduction

1. Introduction

Ramulus Cinnamomi as the commonly used traditional Chinese medicine has the function of calm alleviation pyretic anti-convulsant anti-inflammatory antibacterial diuretic. The main effect of Ramulus Cinnamomi is pungent-warm diaphoresis in traditional Chinese medicine. Traditional experience thoughts Ramulus Cinnamomi as little brown scent perfect. The identification of modern professional mainly based on characters, microscopic identification, chemical reaction, thin layer chromatography, high performance liquid chromatography to identify the quality of Ramulus Cinnamomi. Now there are plenty of inferior and confuse in the market. It is difficult to evaluate Ramulus Cinnamomi with modern methods. The method this paper supports is easy stable well reproducibility master easily

2. The Characteristic of Rough Set

The characteristic of Rough Set [1] is as follows

1) The rough set needn't experience knowledge. Fuzzy set and probability stat is the common method of deal with uncertain information but these methods need some data additional information and experience knowledge. For example: the membership function and probability distributing. Sometimes this information doesn't get easily. The analysis of rough set only make use of the information who support itself needn't any experience knowledge.

2) Rough set is a powerful analysis data facility. It can express and deal with the imperfection information and simplify data and obtain the smallest formula of knowledge and identify and evaluate the rely relationship and reveal the simple concept pattern. It can also obtain the regulation knowledge from experience data especially used intellective control.

In the theory of rough set knowledge can be thought a ability of classifying the realism and abstract object. If there is some knowledge about field and describe the object in the field by property and its value. For

[+] Corresponding author.
E-mail address: linjianyangcmu@126.com.

example the space object assemble U has two property color and shape. The property of color chooses red yellow and green, the property of shape choose square circle triangle. Color and shape compose a group of equal relationship on U from the perspective of disperse math. According to the equal relationship of color the object on U can divide into the assembly red ones yellow ones and green ones. According to the relationship of shape can divide into the assembly square ones circle ones and triangle ones. According to the compound equal relationship of color adding shape can divide into the assembly red circle ones yellow square ones green triangle ones. If the two objects all belong to the assembly of red circle ones and their relationship is indistinguishable because the property described is both red and yellow. The concept of indistinguishable relationship is the basic of rough set theory. It reveals granular structure of the field knowledge.

3. Adaptive Weight PSO

The step of adaptive weight PSO [2] is as follows

①the speed and position of every atom in random initialize series.

②evaluate the fitness of every atom and store the current each atom's position and fitness into each atom's pbest and store the best fitness individual position and fitness into gbest.

③renew the speed and position with the formula as follows

$$v_{i,j}(t+1) = w v_{i,j}(t) + c_1 r_1 [p_{i,j} - x_{i,j}(t)] + c_2 r_2 [p_{g,j} - x_{i,j}(t)]$$

$$x_{i,j}(t+1) = x_{i,j}(t) + v_{i,j}(t+1), \quad j = 1, 2, \ldots, d$$

④renew weght

$$w = \begin{cases} w_{\min} - \dfrac{(w_{\max} - w_{\min}) * (f - f_{\min})}{f_{avg} - f_{\min}}, & f \le f_{avg} \\ w_{\max}, & f > f_{avg} \end{cases}$$

In it w_{max}、w_{min} is the max and min of w respectively f is the current aim function of the atom f_{avg}和 f_{min} is the average aim and the min aim of all atoms respectively. Adaptive weight changes automatically with the aim function of atom in upon formula. So it called adaptive weight.

When the aim of every atom tend to accordant or partly perfect increase the inertia weight but when the aim of atom tend to decentralization decrease the inertia weight. At the same time the corresponding inertia weight gene of atom whose aim function tend to average aim is more less then protect itself. Then protect this atom if not the atom whose aim function worse than average aim function close to better searching area because its inertia weight is larger.

⑤for every particle compare fitness with its best experienced position. If it is better then put it as the recently best position compare all the value of pbest with gbest then renew gbest.

⑥if satisfy the finish condition (usually as the supposed operation precision or iterative degree) then stop searching then output the result or return (③)go on searching.

4. Powell Optimization

POWELL optimization [3] is described as follows:

(1)Select the initial point $X^{(o)}$ and The direction of a set of linearly independent $e^{(i)}$ (i=1,2,⋯,N), Usually $e^{(i)}$ is the N-axis direction, That is, $e^{(i)}=(0,0,\cdots,1,0,\cdots,0)^T$ (i=1,2,⋯,N).

(2)To conduct N times one-dimensional search from the $X^{(o)}$ along the direction of $e^{(i)}$ (i=1,2,⋯,N), get

$$X^{(i)} = X^{(i-1)} + \lambda_i e^{(i)}, \quad \text{i=1,2,⋯,N}$$

$$f(X^{(i)}) = \min_{\lambda} f(X^{(i-1)} + \lambda_i e^{(i)}), \quad \text{i=1,2,⋯,N}$$

Upon completion of it, get $X^{(N)}$. Here λ_k is a constant which is supposed by Characteristics of a problem.

(3) Calculate the change of the function which is at the most rapid ascendant function.

$$DEL \equiv \max_{1\le i\le N}\left|f(X^{(i)}-f(X^{(0)})\right| = \left|f(X^{(ibig)}-f(X^{(0)})\right|$$

(4) Introduce the direction

$$e = X^{(N)} - X^{(0)}, \qquad PTT \equiv 2X^{(N)} - X^{(0)}$$

Calculate

$$f_E \equiv f(PPT) = f(2X^{(N)} - X^{(0)})$$

(5) If $f_E \ge f_0$ Or when $f_E < f_0$ and

$$2(f_0 - 2f_N + f_E)[(f_0 - f_N) - DEL]^2 \ge (f_0 - f_E)^2 DEL$$

Make $X^{(N)}$ as a new starting point, Along a group of old direction above $e^{(i)}$ (i=1,2,⋯,N), to repeat the steps above, go to program(2).

(6) If the conditions of (5)are not met, Along the direction of $e = X^{(N)} - X^{(0)}$, making $X^{(N)}$ as the starting point, to conduct search to get the Minimum point P of the objective function at this direction. Remove $e^{(ibig)}$ the original direction of $e^{(ibig)}$, reserve original N-1 directions, plus other N directions: $e^{(1)}, e^{(2)}, \cdots, e^{(N)}$, put the P of this moment as the starting point,repeat the steps above,that is go to program (2).

(7) If

$2\left|f(X^{(N)}) - f(X^{(0)})\right| / (\left|f(X^{(N)})\right| + \left|f(X^{(0)})\right|) < \varepsilon$ (ε is the accuracy) stop calculating.

5. Normal-Weibull Similarity

It will need Normal-Weibull similarity when calculate corresponding top. According to distributing theory [4] can easily obtain normal-weibull similarity expression.

Simple data y are normal distribution and normal data x are weibull distribution, the probability density function are

$$g(y) = \frac{1}{\sqrt{2\pi}\sigma_y}\exp\left[-\frac{(y-\mu_y)^2}{2\sigma_y^2}\right]$$

$$(-\infty < y < \infty)$$

where μ_y is mean of forces; σ_y is standard deviation of forces.

$$f(x) = \frac{\beta}{\eta}\left(\frac{x-\gamma}{\eta}\right)^{\beta-1}\exp\left[-\left(\frac{x-\gamma}{\eta}\right)^{\beta}\right] \qquad \gamma \le x$$

where η is scale parameters; β is shape parameters; γ .is location parameters.

Order z=|x-y|, by integral method of the probability density function, probability density function of z can be gotten

$$h(z) = \int_{-\infty}^{0} f(z+y)g(y)dy = \frac{1}{\sqrt{2\pi}\sigma_y}\cdot$$

$$\int_{-\infty}^{0}\frac{\beta}{\eta}\left(\frac{z+y-\gamma}{\eta}\right)^{\beta-1}\exp\left[-\left(\frac{z+y-\gamma}{\eta}\right)^{\beta}\right]\cdot\exp\left[-\frac{(y-\mu_y)^2}{2\sigma_y^2}\right]dy$$

The similarity formula is

$$R_1 = \int_{-\infty}^{0} h(z)dz$$

Similarity ratio

It need similarity ratio [5] when calculate the total similar and its expression is

$$C_{ir} = \frac{\sum_{k=1}^{m} X_{ik} \cdot X_{rk}}{\sum_{k=1}^{m} X_{ik}^2 + \sum_{k=1}^{m} X_{rk}^2 - \sum_{k=1}^{m} X_{ik}^2 \sum_{k=1}^{m} X_{rk}^2}$$

In it X_{ik} is the i sample and k diagnostic variable(k=1,2 ⋯ m) and X_{rk} is the k character variable of common pattern average vector(k=1,2 ⋯ m)

6. Integral and Interpolation

This paper use lobalo integral and Hermite interpolation [6].

The lobato quadrature formula is as follows

$$I(f) = \int_{-1}^{1} f(x)dx \approx \frac{2}{n(n-1)} \cdot$$
$$[f(-1) + f(1)] + \sum_{k=2}^{n-1} A_k f(x_k)$$

In it f is integral function, n is the number of integral points, A_K is user-defined coefficient and X_K is user-defined integral point. The node and coefficient of gauss-lobato quadrature formula is in the table 1.

TABLE I. THE NODE AND COEFFICIENT OF LOBATO FORMULA

n	X_k	A_k
3	0	1.333333
	±1	0.333333
4	±0.447214	0.833333
	±1	0.166666
5	0	0.711111
	±0.654654	0.544444
	±1	0.100000
6	±0.285232	0.554858
	±0.765055	0.378475
	±1	0.066667

Hermite interpolation

Hermite interpolation satisfies that it is equal to the given function in the node and the coefficient in it is equal to the given coefficient. Hermite interpolation multinomial is much complex to the condition of high-level coefficient. There always the function and one-level coefficient has been given in actually. In this condition the expression of n nodes Hermite interpolation multinomial H(x) is as follows.

$$H(x) = \sum_{1}^{n} h_i[(x_i - x)(2a_i y_i - y_i) + y_i]$$

In it

$$y_i = y(x_i),$$
$$h_i = \prod_{\substack{j=1 \\ j \neq i}}^{n} \left(\frac{x - x_j}{x_i - x_j} \right)^2, a_i = \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{1}{x_i - x_j}$$

7. Ramulus Cinnamomi Quality Assessment

The original test data[7] of Ramulus Cinnamomi quality are in table 2, the data in the table is the comparative chromatogram top acreage of Ramulus Cinnamomi，the first line is the chromatogram number after reduce distinguish function and the first row is the criterion number of different producing area. In it the tenth sample is shoddy goods and the data don't delete and participate in the account.

Table 1 The original data of Ramulus Cinnamomi quality

No	Xl	X2	X3	X4	X5	X6	X7	X8	X9	X10
1	0.00	4.80	0.78	0.10	0.10	11.3	0.00	546.9	0.0	28.4
2	0.00	1.76	5.71	0.06	0.17	9.3	0.00	474.8	0.0	22.6
3	0.00	0.00	2.40	0.00	2.65	5.1	0.00	191.5	12.9	2.9
4	0.06	1.00	1.93	0.00	0.14	1.5	0.00	730.2	0.0	7.7
5	0.00	3.33	7.49	0.01	4.77	1.8	0.00	432.5	10.6	25.5
6	0.00	0.97	6.93	0.00	0.40	6.0	0.00	356.8	0.0	28.6
7	0.07	5.30	8.94	0.12	0.11	1.3	0.00	558.4	0.0	36.3
8	0.00	3.16	8.48	0.13	0.25	2.0	0.00	434.8	27.5	24.4
9	0.26	5.84	24.86	0.00	0.10	1.4	1.55	396.6	10.4	65.7
11	0.05	4.81	10.96	0.23	4.25	2.2	0.00	389.9	0.3	27.8
12	0.16	2.42	8.23	0.07	0.10	9.5	1.02	329.3	0.0	25.5
13	0.11	2.96	14.40	0.12	0.78	26.8	0.00	449.3	5.9	30.4
14	0.19	5.34	12.02	0.37	0.93	29.6	1.54	480.6	0.9	43.0
15	0.16	2.05	10.29	0.20	0.23	16.6	0.87	298.0	0.5	26.7
16	0.00	0.00	6.00	0.00	0.10	10.2	0.00	199.3	4.4	7.3
17	0.00	0.00	3.94	0.13	0.07	6.6	0.00	182.2	3.4	5.5
18	0.39	2.96	8.68	0.15	0.35	22.6	1.23	404.9	6.4	18.4
19	0.18	1.16	6.56	0.07	0.48	19.6	0.64	310.6	1.0	18.2
20	0.15	1.92	5.92	0.06	0.69	23.4	0.42	381.6	1.9	19.6
21	0.00	0.97	5.98	0.12	0.89	15.9	0.00	295.4	1.8	13.8
22	0.29	2.83	5.72	0.00	1.74	12.0	0.57	387.0	3.7	15.6
23	0.45	1.00	7.45	0.12	0.23	13.7	0.55	397.3	2.7	18.3
24	0.04	3.42	2.21	0.09	0.15	16.1	0.00	298.4	0.5	8.0
25	0.11	3.10	13.81	0.09	0.10	13.9	0.00	369.8	0.8	34.0
26	0.00	2.28	8.42	0.00	0.00	21.9	0.81	385.6	1.8	28.1
27	0.17	2.94	12.10	0.43	0.12	31.5	2.20	400.7	1.8	51.0
28	0.11	6.63	10.97	0.27	0.23	19.7	0.00	565.6	1.0	39.7
29	0.12	5.75	7.59	0.17	0.31	13.1	0.00	531.0	1.2	46.3
30	0.09	2.17	8.91	0.11	0.00	17.7	0.00	387.7	1.1	27.2
31	0.00	0.82	3.02	0.07	0.00	8.2	0.00	153.4	0.8	8.9
32	0.00	0.46	3.19	0.00	0.00	13.0	0.00	357.6	1.6	10.5
33	0.00	0.23	4.04	0.00	0.00	8.7	0.00	155.7	0.9	9.1
34	0.00	0.51	4.41	0.00	0.30	14.2	0.00	176.4	1.9	8.8
35	0.00	0.37	2.77	0.00	0.00	5.6	0.00	162.4	1.2	16.0
36	0.00	0.36	7.88	0.00	0.00	17.5	0.00	226.6	3.3	18.3
37	0.00	0.41	4.60	0.00	0.00	13.3	0.00	254.4	2.7	12.7
38	0.00	1.93	6.87	0.10	0.22	13.7	0.00	355.4	5.1	16.4
39	0.08	0.25	9.36	0.12	0.17	13.1	0.00	396.8	3.6	31.0
40	0.00	0.40	1.78	0.27	0.46	12.9	0.00	170.9	3.6	8.4
41	0.00	0.77	2.33	0.10	2.69	7.3	0.00	256.4	0.8	5.1
42	0.00	1.18	3.87	0.00	0.13	9.5	0.00	292.9	0.4	12.6
43	0.00	1.30	4.84	0.00	0.14	18.9	0.00	184.9	0.6	12.7

Data are clustered and divided into three discrete levels, the results in table 3

Table 2 The data clustered and discretizated

No	Xl	X2	X3	X4	X5	X6	X7	X8	X9	X10	D
1	2	1	2	3	2	3	2	1	1	3	1
2	2	2	3	3	2	1	2	1	1	3	1
3	2	3	2	2	3	1	2	3	2	1	2
4	2	3	2	2	2	1	2	1	1	1	1
5	2	2	3	2	3	1	2	2	2	3	1
6	2	3	3	2	2	1	2	2	1	3	1
7	2	1	3	3	2	1	2	1	1	3	1
8	2	2	3	3	2	1	2	2	2	3	1
9	3	1	1	2	2	1	1	2	2	2	1
11	2	1	3	1	3	1	2	2	1	3	1
12	1	2	3	3	2	1	1	2	1	3	1
13	1	2	1	3	1	2	2	2	3	3	1
14	1	1	1	1	1	2	1	1	1	2	1
15	1	2	3	3	2	3	1	3	1	3	2
16	2	3	3	2	2	1	2	3	3	1	2
17	2	3	2	3	2	1	2	3	3	1	2
18	3	2	3	3	2	2	1	2	3	3	1
19	1	3	3	3	2	3	2	2	1	3	2
20	1	2	3	3	1	2	2	2	1	3	1
21	2	3	3	3	1	3	2	3	1	1	2
22	3	2	3	2	1	3	2	2	3	1	1
23	3	3	3	3	2	3	2	2	3	3	1
24	2	2	2	3	2	3	2	3	1	1	2
25	1	2	1	3	2	3	2	2	1	3	1
26	2	2	3	2	2	2	1	2	1	3	2
27	1	2	1	1	2	2	1	2	1	2	1

28	1	1	3	1	2	3	2	1	1	2	1
29	1	1	3	3	2	3	2	1	1	2	1
30	1	2	3	3	2	3	2	2	1	3	1
31	2	3	2	3	2	1	2	3	1	1	2
32	2	3	2	2	2	3	2	2	1	1	1
33	2	3	2	2	2	1	2	3	1	1	2
34	2	3	2	2	2	3	2	3	1	1	2
35	2	3	2	2	2	1	2	3	1	1	2
36	2	3	3	2	2	3	2	3	3	3	2
37	2	3	2	2	2	3	2	3	3	1	2
38	2	2	3	3	2	3	2	2	3	1	1
39	1	3	3	3	2	3	2	2	3	3	1
40	2	3	2	1	2	3	2	3	3	1	2
41	2	3	2	3	3	1	2	3	1	1	2
42	2	3	2	2	2	1	2	3	1	1	2
43	2	3	2	2	2	3	2	3	1	1	2

Reduce attributes to data of Table 3 with matlab rough tools [8] then we can get the core set {C2, C8, C9} induced from C attributes selected to D. In it the centre data of stair clustering is 3.0511±0.01526、439.41±2.19705、3.6663±0.01833 and the centre data of secondary clustering is 0.9132±0.00457、227.643±1.13822、2.134±0.01067。

8. Conclution

1) It can avoid the partly perfect with adaptive weight PSO and it can get the centre actual data of Ramulus Cinnamomi every clustering with POWELL optimization.

2) It can deflate the criterion data-base of Ramulus Cinnamomi with knowledge reduction and decrease the complex of arithmetic 、increase the identification effect of the quality. Because eliminate unnecessary interfere make the identification process more clear.

3) It can ensure account simple and resemble account precision and avoid the magnify of accumulate error and boost constringency speed at the same time with lobato integral and Hermite interpolation.

9. Acknowledgment

This work is supported by PHD Research Start Funds of Liaoning Province under Grant 20091111, Science Research Project of Liaoning Province Education Office under Grant 2009A721.

10. References

[1] B Ji. Research on rough set theory and its progresson application [J].Computer Technology and Development, 2007, 17(3):69-72.

[2] C Gong, Z L Wang. Proficient optimization calculation in Matlab [M]. Beijing: Publishing house of electronics industry, 2009:283.

[3] G Y He. VB commonly used algorithm daquan [M]. Xi'an:Xi'an electronic technology university press, 2001: 340-341.

[4] Y M Zhang. Reliable design of vehicle spare-parts [M].Beijing: Beijing Institute of Technology Press, 2000, 71-74.

[5] X Wang, W Y Wang, K R Zhang, etc. Aproaching the study on the similarity analysis of HPLC fingerprints spectra china medinicens [J]. Journal of Shenyang Pharmaceutica, 2003,20(5): 361-362.

[6] C Gong，Z L Wang. MATLAB assembly language commonly used algorithm [M]. Beijing: Publishing House of Electronics Industry, 2008: 84, 198-199.

[7] S Yang. Study on the Method of Quality Assessment of Ramulus Cinnamomi [D]. Shenyang: Shenyang pharmaceutical university,2002:45-46.

[8] D Y Xue, Y Q Chen. Matlab solution for the advanced mathematic program [M].Beijing: Tsinghua，2008:365-368.