Applied Soft Computing 25 (2014) 514–518

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Short Communication

Fuzzy fractal dimension of complex networks

Haixin Zhang[a,b], Yong Hu[c], Xin Lan[a], Sankaran Mahadevan[d], Yong Deng[a,d,*]

[a] School of Computer and Information Science, Southwest University, Chongqing 400715, China
[b] College of Computer Science, Chongqing University, Chongqing 400044, China
[c] Institute of Business Intelligence and Knowledge Discovery, Guangdong University of Foreign Studies, Guangzhou 510006, China
[d] School of Engineering, Vanderbilt University, Nashville, TN 37235, USA

ARTICLE INFO

Article history:
Received 17 March 2013
Received in revised form 21 February 2014
Accepted 11 August 2014
Available online 3 September 2014

Keywords:
Fractal dimension
Complex network
Box-covering dimension
Fuzzy sets

ABSTRACT

Complex networks are widely used to describe the structure of many complex systems in nature and society. The box-covering algorithm is widely applied to calculate the fractal dimension, which plays an important role in complex networks. However, there are two open issues in the existing box-covering algorithms. On the one hand, to identify the minimum boxes for any given size belongs to a family of Non-deterministic Polynomial-time hard problems. On the other hand, there exists randomness. In this paper, a fuzzy fractal dimension model of complex networks with fuzzy sets is proposed. The results are illustrated to show that the proposed model is efficient and less time consuming.

1. Introduction

Complex network plays more and more important role in academic researches [1–5]. It has been shown that the small-world property [6] and the scale-free property [7] are the two fundamental properties of complex networks. The small-world property reveals that the number of nodes of the complex network, denoted by N, increase exponentially with the average diameter of the network, denoted by $\bar{l}$, which leads to the general understanding that complex networks are not self-similar, since self-similar requires a power-law relation between N and $\bar{l}$. Recently, Song et al. found that a variety of real complex networks consist of self-repeating patterns on all length scales using the box-covering algorithm [8,9]. From then on, fractal and self-similarity properties of complex networks attract much attention from many researchers [10–14,5,15].

The box-covering algorithm is widely applied to calculate the fractal dimension of complex networks [9,16–18,13]. The main step in box-covering algorithm is to find the minimum number of boxes of given size to cover the nodes of the complex network. The fractal dimensions of complex networks is determined by the relationship between the minimum number of boxes and the size of the box. However, there are two main problems during the process to find the minimum number of boxes. On the one hand, how to cover a network with the fewest possible number of boxes of a given size belongs to a family of Non-deterministic Polynomial-time hard (NP-hard) problems [19], which means only approximate solution can be found by an algorithm in an acceptable amount of time. On the other hand, there exists randomness [9], such as randomly ranking the nodes in a sequence or randomly selecting the node as the center, which means their results fluctuate randomly. In other words, it is necessary to apply an average process to obtain a reliable result. The averaging process needs a large amount of time consuming.

Fuzzy theory is an effective tool in many fields [20–27], like to deal with uncertain information [28–33] and fractal dimension model [34]. However, best to our knowledge, the fuzzy sets is not applied in complex network. To address the open issues mentioned before. In this paper, a fuzzy fractal dimension model of complex networks with fuzzy sets is proposed. The results are shown that the proposed fuzzy fractal dimension model is efficient and less time consuming.

2. Preliminaries

In this section, some basic concepts of complex network, box-covering algorithm for fractal dimension of complex network and fuzzy sets are briefly introduced.

* Corresponding author at: School of Computer and Information Science, Southwest University, Chongqing 400715, China. Fax: +86 023 68254555.
E-mail addresses: ydeng@swu.edu.cn, professordeng@163.com, prof.deng@hotmail.com (Y. Deng).

http://dx.doi.org/10.1016/j.asoc.2014.08.019
1568-4946/

2.1. Complex network

For complex network $G=(N, V)$ and $N=(1, 2, \ldots, n)$, $V=(1, 2, \ldots, m)$, where n is the total number of nodes, m is the total number of edges. G is defined as unweighted network when the cell $x_{ij}(i, j=1, 2, \ldots, n)$ of edge is defined as 1 if node i is connected to node j, and 0 otherwise.

Shortest path is measure of finding a path between two nodes in the complex network such that the sum of values of its constitute edge is minimized [35]. For the unweighted networks, shortest path length is defined as follows.

Definition 2.1 *(the shortest path length)*. Denote d_{ij} as the shortest path length between node i and node j, which satisfies

$$d_{ij} = \min(x_{ih} + \ldots + x_{hj}) \tag{1}$$

The diameter of a network is defined as:

Definition 2.2 *(the diameter of a networks)*. Let d_{ij} be the shortest path between node i and node j of network $G=(N, V)$. Then the diameter of G, denoted by d, is

$$d = \max(d_{ij}) \tag{2}$$

2.2. Box-covering dimension of complex network

The original definition of box-covering is initially proposed by Hausdorff [36,37]. It is initially applied in the complex networks by Song et al. [8,9]. For a given box size l_B, a box is a set of nodes where all distances d_{ij} between any two nodes i and j in the box are smaller than l_B. The minimum number of boxes is denote by N_B, which must be covered the entire network. Finally, the fractal dimension of this network is denoted by d_B and given as follows:

$$d(B) \approx -\frac{\ln N_B}{\ln l_B} \tag{3}$$

The idea of box-covering algorithm of the complex networks has three steps and shown in Fig. 1 [9].

Step 1: For given network G_1 and box size l_B. A new network G_2 is obtained, in which node i is connected to node j when $d_{ij} \geq l_B$.
Step 2: Greedy coloring algorithm is adopted to find the minimum number of colors to color the nodes of network G_2 in which the color of each node is different from its nearest neighbors' colors. Then a colored network G_3 is obtained.
Step 3: Let one color represent one box in network G_4. Then, we obtain the value of N_B, equals to the number of different colors.

The problem of finding the minimum number of boxes in G_1 was translated to a graph coloring problem in G_2 in Song's box-covering algorithm. Graph coloring problem is a well-known NP-hard problem, which means only approximate solution can be found by an algorithm in an acceptable amount of time. In addition, the results of Song's box-covering algorithm depend on the original sequence of the nodes of the network. So an averaging process was needed. For more detailed information please refer to [9].

2.3. Fuzzy sets

In traditional two-state classifiers, where a class C is defined as a subset of the universal set X, any input pattern $x \in X$ can either be a member or not be a member of the given class C. This property of whether or not a pattern x of the universal set belongs to the class C can be defined by a characteristic function $\mu_C: X \to \{0, 1\}$ as follows: $\mu_C(x)=1$ iff $x \in C$; $\mu_C(x)=0$ otherwise. In real life situations, boundaries between the classes may be overlapping. Hence, it is uncertain whether an input pattern belongs totally to the class C. To consider such situations, in fuzzy sets the concept of the characteristic function is modified to the fuzzy membership function μ_C: $X \to [0, 1]$.

3. Fuzzy fractal dimension model of complex networks

The box-covering dimension of complex networks employed the original definition of box-covering by Hausdorff [10,13]. The fractal dimension d_B follows the relation

$$N(\varepsilon) \approx \varepsilon^{-d_B} \tag{4}$$

where $N(\varepsilon)$ is the minimum number of boxes needed to cover a given network, and ε is the box size. The box of size ε is a set of nodes where the distance between any pair of nodes is less than

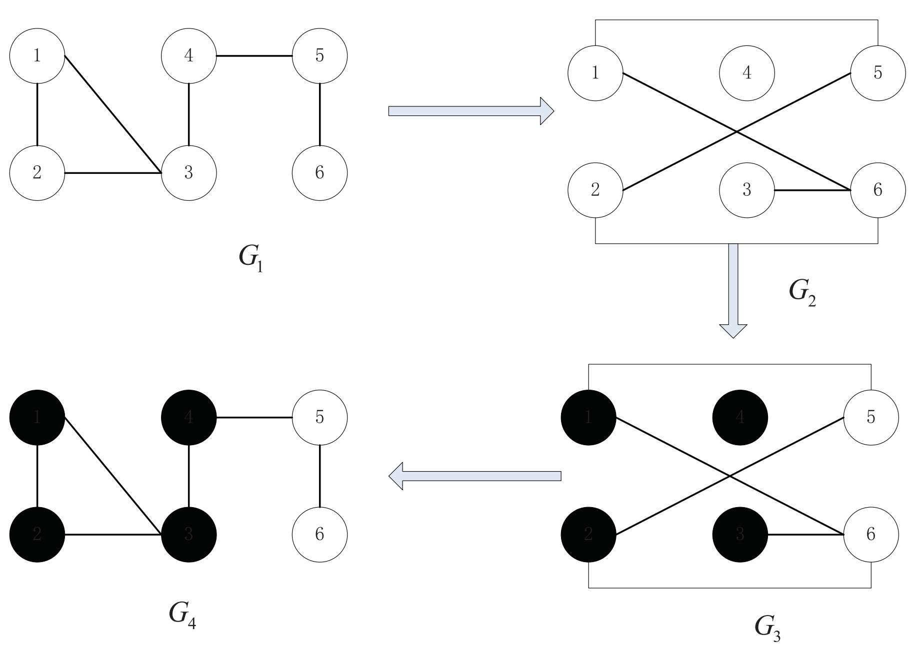

Fig. 1. The idea of the Song's box-covering algorithm for the unweighted networks, where $l_B=3$.

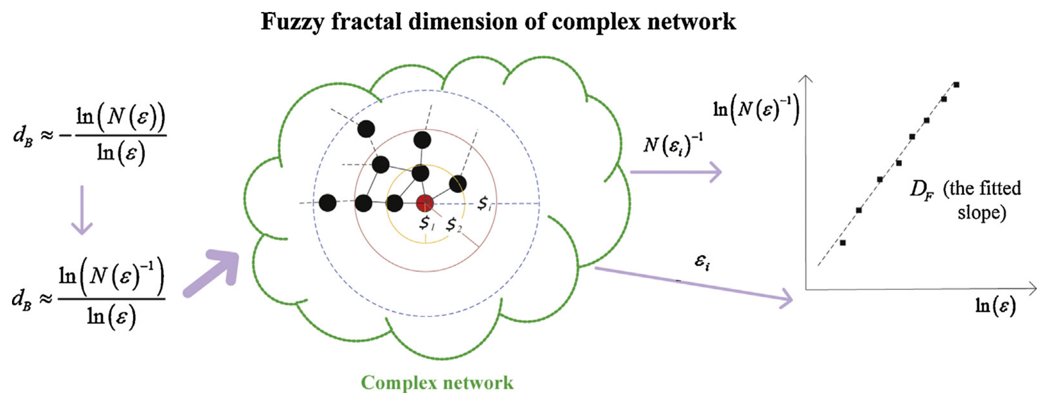

Fig. 2. The model of the fuzzy fractal dimension of complex network. Calculate the covering ability (CA) $N(\varepsilon_i)^{-1}$ of a ball of radius ε_i. Regress $\ln(N(\varepsilon)^{-1})$ vs. $\ln(\varepsilon)$. The fuzzy fractal dimension D_F can be the fitted slope.

ε [10], or a set of nodes whose shortest distance from the center v_c is less than ε [13]. In practice, the fractal dimension d_B can be calculated by the following equation:

$$d_B \approx -\frac{\ln(N(\varepsilon))}{\ln(\varepsilon)} \tag{5}$$

The equation can be rewritten as:

$$d_B \approx \frac{\ln(N(\varepsilon)^{-1})}{\ln(\varepsilon)} \tag{6}$$

As mentioned before, $N(\varepsilon)$ is the minimum number of boxes needed to cover a given network. $(N(\varepsilon))^{-1}$ is the reciprocal of $N(\varepsilon)$. $(N(\varepsilon))^{-1}$ can be seen as the covering ability (CA) of boxes. Large value of $N(\varepsilon)$ corresponds to small value of $(N(\varepsilon))^{-1}$. That is to say, the more boxes are needed to cover the network, the less percentage of nodes of the network could be covered by a box. For example, if five boxes are needed to cover the network, then one box can cover one fifth nodes of the network. According to the idea of box covering ability, the fuzzy fractal dimension model is proposed. The general process is shown in Fig. 2.

The main step in box-covering algorithm is to find the minimum number of boxes of given size to cover the nodes of the complex network, while the main step of our model is to calculate the box covering ability. Identical balls of radius ε are constructed around every node of the complex network, $G = (N, V)$. The covering ability (CA) of a ball of radius ε for this complex network is calculated by the following equation:

$$N(\varepsilon)^{-1} = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} \Omega_{ij}(\varepsilon) A_{ij}(\varepsilon) \tag{7}$$

where

$$\Omega_{ij}(\varepsilon) = \begin{cases} 1, & d_{ij} \leq \varepsilon \\ 0, & otherwise \end{cases} \tag{8}$$

and

$$A_{ij}(\varepsilon) = \exp\left(-\frac{d_{ij}^2}{\varepsilon^2}\right) \tag{9}$$

d_{ij} is the shortest path between node i and node j. $\Omega_{ij}(\varepsilon)$ is the selecting function, 1 represents that the node j could be covered by the ball whose center is node i, 0 otherwise. $A_{ij}(\varepsilon)$ is a membership function, with the value change from 0 to 1, which represents the degree of how node j could be covered by the ball. The less distance from node j to the ball center node i, the higher the value is.

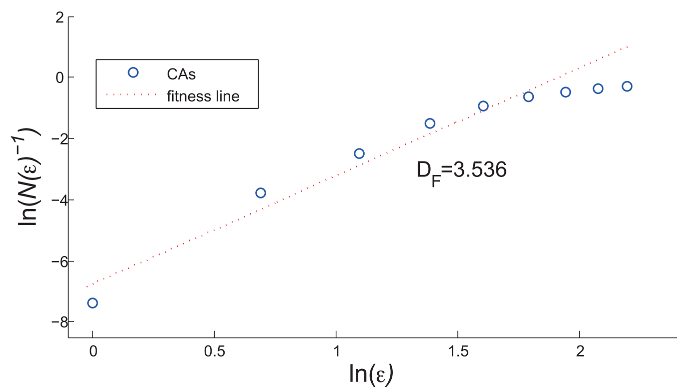

Fig. 3. The fuzzy fractal dimension D_F of *E. coli* network.

For instance, in G_1 in Fig. 1, the CA of a ball of size of radius 3 centered by node 3 is calculated as:

$$\begin{aligned} N_3(3)^{-1} &= \exp\left(-\frac{d_{31}^2}{3^2}\right) + \exp\left(-\frac{d_{32}^2}{3^2}\right) + \exp\left(-\frac{d_{34}^2}{3^2}\right) \\ &\quad + \exp\left(-\frac{d_{35}^2}{3^2}\right) + \exp\left(-\frac{d_{36}^2}{3^2}\right) \\ &= \exp\left(-\frac{1^2}{3^2}\right) + \exp\left(-\frac{1^2}{3^2}\right) + \exp\left(-\frac{1^2}{3^2}\right) + \exp\left(-\frac{2^2}{3^2}\right) + \exp\left(-\frac{3^2}{3^2}\right) \\ &= 0.7387 \end{aligned}$$

we obtained the CAs of other nodes in a same manner as follows

$$N_1(3)^{-1} = 0.5936$$

$$N_2(3)^{-1} = 0.5936$$

$$N_4(3)^{-1} = 0.7426$$

$$N_5(3)^{-1} = 0.6333$$

$$N_6(3)^{-1} = 0.4484$$

Averaging these CAs, we get the CA of a ball of size of radius 3 for this complex network is $N(3)^{-1} = 0.6250$, which means a ball of size of radius 3 could cover 62.5% nodes of the network. In other words, $1/0.6250 = 1.6000$ boxes are needed to cover the network. This is an average number in different cases. If the center is node 3, then, only one box is needed, if the center is node 1 or node 2, then, another box is needed to cover node 6, and so on. To get the CA is a smoothing process, the complexity reduced significantly from NP-hard problems in box-covering algorithm.

Table 1
Results comparison between our method and Song's method.

	Fractal dimension	Time complexity	Time consuming
Proposed method	3.5358	$O(dn^2)$	$O(dn^2)$
Song's method	3.47 ± 0.11	$O(dn^2)$	$10,000 * O(dn^2)$

Change the size of radius ε from 1 to half of the diameter d of the network, the corresponding covering ability (CA) of a ball of such size could be obtained. Then the fuzzy fractal dimension D_F of this complex network can be calculated from

$$D_F \approx \frac{\ln(N(\varepsilon)^{-1})}{\ln(\varepsilon)} \tag{10}$$

The algorithm to compute the fuzzy fractal dimension D_F of complex networks is illustrated in Algorithm 1.

Algorithm 1. Compute the fuzzy fractal dimension of complex network.

Input: A complex network G of N nodes.
  for each ε in$[\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_L](L = d/2)$ **do**
    $N(\varepsilon)^{-1} = 0$
    for each i in[1,N] **do**
      for each j in[1,N] **do**
        find the shortest path length d_{ij} between node N_i and node N_j
        if $d_{ij} \le \varepsilon$ **then**
          $A_{ij} = \exp\left(-\frac{d_{ij}^2}{\varepsilon^2}\right)$
          $N(\varepsilon)^{-1} = N(\varepsilon)^{-1} + A_{ij}$
        end if
      end for
    end for
    $N(\varepsilon)^{-1} = \frac{N(\varepsilon)^{-1}}{N(N-1)}$
  end for
  regress $\ln(N(\varepsilon)^{-1})$ vs. $\ln(\varepsilon)$
  fit the slope of the regressed line, D_F
Output: The fuzzy fractal dimension D_F of complex network G.

In our algorithm, only one value of the fractal dimension could be obtained and only need one experiment. In the existing box-covering algorithm [9], the results of the greedy algorithm may depend on the original coloring sequence. So, the value of the fractal dimension also depends on the original coloring sequence. In order to investigate the quality of the algorithm, the coloring sequence should be randomly reshuffled and the greedy algorithm should be applied for many times, so it takes a lot of time.

4. Results and discussion

In this section, the *Escherichia coli* network [13], with 2859 proteins and 6890 interactions between them, was investigated. The results from our model is shown in Fig. 3.

First, calculate the CAs for different value of radius by Eqs. (7)–(9). Then, regress $\ln(N(\varepsilon)^{-1})$ vs. $\ln(\varepsilon)$. It is shown that the covering ability $N(\varepsilon)^{-1}$ and ball radius ε follows a power-low relationship. Finally, the fitness line is obtained so as to capture the fuzzy fractal dimension by Eq. (10). The fractal dimension of *E. coli* network is 3.5358.

The results and the time consuming of our algorithm and Song's box-covering algorithm were compared in Table 1. From the results, we can see that the fractal dimension D_F through our model is 3.5358, which is consistent with the range 3.47 ± 0.11 from Song's box-covering algorithm [13]. The time complexity of both algorithms are $O(dn^2)$. However, due to the randomness of the sequence of the nodes in Song's box-covering algorithm, the fractal dimension can not be obtained by one experiment, but be averaged for a lot of experiments. Song et al. themselves took 10,000 times in their work [9], so the time consuming of Song's box-covering algorithm is 10,000 times increased in practice.

In Song's box-covering algorithm, the value of the fractal dimension depends on the original coloring sequence. The results are different in different routine experiments for the same complex network. It is not coincide with our intuition. Given a certain complex network, we believe that its fractal dimensional is also a certain value. Our model could provide a deterministic result. Moreover, our model can obtain the result only through one experiment, while box-covering algorithms need an averaging process, which needs a large number of experiments, so it takes a lot of time. In addition, in Song's box-covering algorithm, a matrix of the number of the nodes of the network rows and the value of the diameter of the network columns [9] is required. The matrix is a big challenge to the computer when the network is very large, like a network of one million nodes.

5. Conclusion

A fuzzy fractal dimension for complex networks with fuzzy sets is proposed in this paper. This model could provide a deterministic fractal dimension for a certain network, while in box-covering algorithms, the fractal dimensional values fluctuate randomly even facing the same complex network. Our model can provide the result only through one experiment, while box-covering algorithms need a large number of times of experiments for averaging, which is time consuming. The results are illustrated to show that the proposed model is easy and useful.

Acknowledgements

The author greatly appreciate the reviews' suggestions. The work is partially supported by National Natural Science Foundation of China (Grant Nos. 61174022 and 71271061), R&D Program of China (2012BAH07B01), National High Technology Research and Development Program of China (863 Program) (Grant No. 2013AA013801), Science and Technology Planning Project of Guangdong Province, China (Grant Nos. 2010B010600034, 2012B091100192), China State Key Laboratory of Virtual Reality Technology and Systems.

References

[1] R. Albert, A.-L. Barabási, Statistical mechanics of complex networks, Rev. Mod. Phys. 74 (2002) 47–97.
[2] M.E. Newman, The structure and function of complex networks, SIAM Rev. 45 (2003) 167–256.
[3] C. Shi, Z. Yan, Y. Cai, B. Wu, Multi-objective community detection in complex networks, Appl. Soft Comput. 12 (2012) 850–859.
[4] U. Manzoor, S. Nefti, iDetect: content based monitoring of complex networks using mobile agents, Appl. Soft Comput. 12 (2012) 1607–1619.
[5] D. Wei, X. Deng, X. Zhang, Y. Deng, S. Mahadevan, Identifying influential nodes in weighted networks based on evidence theory, Phys. A: Stat. Mech. Appl. 392 (2013) 2564–2575.
[6] D.J. Watts, S.H. Strogatz, Collective dynamics of small-world networks, Nature 393 (1998) 440–442.
[7] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (1999) 509–512.
[8] C. Song, S. Havlin, H. Makse, Self-similarity of complex networks, Nature 433 (2005) 392–395.
[9] C. Song, L. Gallos, S. Havlin, H. Makse, How to calculate the fractal dimension of a complex network: the box covering algorithm, J. Stat. Mech. Theory Exp. 2007 (2007) P03006.
[10] C. Song, S. Havlin, H. Makse, Origins of fractality in the growth of complex networks, Nat. Phys. 2 (2006) 275–281.
[11] J. Kim, K. Goh, B. Kahng, D. Kim, Fractality and self-similarity in scale-free networks, New J. Phys. 9 (2007) 177.
[12] J. Kim, K.-I. Goh, G. Salvi, E. Oh, B. Kahng, D. Kim, Fractality in complex networks: critical and supercritical skeletons, Phys. Rev. E 75 (2007) 016110.
[13] C. Schneider, T. Kesselring, J. Andrade Jr., H. Herrmann, Box-covering algorithm for fractal dimension of complex networks, Phys. Rev. E 86 (2012) 016707.
[14] C. Fan, J.-L. Guo, Y.-L. Zha, Fractal analysis on human dynamics of library loans, Phys. A: Stat. Mech. Appl. 391 (2012) 6617–6625.
[15] H. Zhang, X. Lan, D. Wei, S. Mahadevan, Y. Deng, Self-similarity in complex networks: from the view of the hub repulsion, Mod. Phys. Lett. B 27 (2013) 1350201.

[16] J. Kim, K.-I. Goh, B. Kahng, D. Kim, A box-covering algorithm for fractal scaling in scale-free networks, Chaos 17 (2007), 026116–026116.
[17] L. Gao, Y. Hu, Z. Di, Accuracy of the ball-covering approach for fractal dimensions of complex networks and a rank-driven algorithm, Phys. Rev. E 78 (2008) 046109.
[18] L. Daqing, K. Kosmidis, A. Bunde, S. Havlin, Dimension of spatially embedded networks, Nat. Phys. 7 (2011) 481–484.
[19] M.R. Garey, D.S. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness, W.H. Freeman & Co., New York, NY, USA, 1990.
[20] C.-C. Chou, L.-J. Liu, S.-F. Huang, J.-M. Yih, T.-C. Han, An evaluation of airline service quality using the fuzzy weighted SERVQUAL method, Appl. Soft Comput. 11 (2011) 2117–2128.
[21] Y. Deng, F.T. Chan, A new fuzzy dempster MCDM method and its application in supplier selection, Expert Syst. Appl. 38 (2011) 9854–9861.
[22] Y. Deng, R. Sadiq, W. Jiang, S. Tesfamariam, Risk analysis in a linguistic environment: a fuzzy evidential reasoning-based approach, Expert Syst. Appl. 38 (2011) 15438–15446.
[23] B. Kang, Y. Deng, R. Sadiq, S. Mahadevan, Evidential cognitive maps, Knowl. Based Syst. 35 (2012) 77–86.
[24] W.-W. Wu, Segmenting critical factors for successful knowledge management implementation using the fuzzy DEMATEL method, Appl. Soft Comput. 12 (2012) 527–535.
[25] Y.-C. Chou, C.-C. Sun, H.-Y. Yen, Evaluating the criteria for human resource for science and technology (HRST) based on an integrated fuzzy AHP and fuzzy DEMATEL approach, Appl. Soft Comput. 12 (2012) 64–71.
[26] X. Zhang, Y. Deng, F.T. Chan, P. Xu, S. Mahadevan, Y. Hu, IFSJSP: a novel methodology for the job-shop scheduling problem based on intuitionistic fuzzy sets, Int. J. Prod. Res. 51 (2013) 5100–5119.
[27] S. Chen, Y. Deng, J. Wu, Fuzzy sensor fusion based on evidence theory and its application, Appl. Artif. Intell. 27 (2013) 235–248.
[28] M.-L. Tseng, Green supply chain management with linguistic preferences and incomplete information, Appl. Soft Comput. 11 (2011) 4894–4903.
[29] M.-L. Tseng, Using a hybrid MCDM model to evaluate firm environmental knowledge management in uncertainty, Appl. Soft Comput. 11 (2011) 1340–1352.
[30] J. Liu, F.T. Chan, Y. Li, Y. Zhang, Y. Deng, A new optimal consensus method with minimum cost in fuzzy group decision, Knowl. Based Syst. 35 (2012) 357–360.
[31] Y. Deng, Y. Chen, Y. Zhang, S. Mahadevan, Fuzzy Dijkstra algorithm for shortest path problem under uncertain environment, Appl. Soft Comput. 12 (2012) 1231–1237.
[32] M.-L. Tseng, Y.-H. Chen, Y. Geng, Integrated model of hot spring service quality perceptions under uncertainty, Appl. Soft Comput. 12 (2012) 2352–2361.
[33] Y. Zhang, Z. Zhang, Y. Deng, S. Mahadevan, A biologically inspired solution for fuzzy shortest path problems, Appl. Soft Comput. 13 (2013) 2356–2363.
[34] W. Pedrycz, A. Bargiela, Fuzzy fractal dimensions and fuzzy modeling, Inf. Sci. 153 (2003) 199–216.
[35] M. Newman, Analysis of weighted networks, Phys. Rev. E 70 (2004) 056131.
[36] H. Jürgens, H. Peitgen, D. Saupe, Chaos and Fractals: New Frontiers of Science, 1992.
[37] A. Bunde, S. Havlin, Fractals in Science: With a MS-DOS Program Diskette, Springer-Verlag, 1994.