Mining Multi-scale Intervention Rules from Time Series and Complex Network[*]

Jiaoling Zheng[1], Changjie Tang[2], Shaojie Qiao[3][†], Ning Yang[2], Yue Wang[2]

[1] *Department of Software Engineering, Chengdu University of Information Technology, Chengdu 610225, China*

[2] *Institute of Database and Knowledge Engineering, School of Computer Science , Sichuan University, Chengdu 610065,China*

[3] *School of Information Science and Technology, Southwest Jiaotong University, Chengdu 610031, China*

Abstract

This paper proposes the concept of intervention rule which tries to reveal the interventional relationship between elements in a system in the following three aspects. (1) Casual relationship. Intervention rule shows which element is the cause and which element is the consequence. (2) Quantitative relationship: Intervention rule shows the quantitative intensity of how the change of the causal element interferes with the change of the consequential element. (3) Multi-scale intervention relationship. Intervention rule shows the intervention at different decomposition scale of the original system, since sub system may exhibit different mechanism from the original system. This paper first introduces a general intervention rule framework, and then transforms the framework into concrete intervention rules for complex network data and time series data. Then, it proposes two algorithms to mine the intervention rules from the two different systems. Finally, the experimental results show that multi-scale intervention rules do exist in real dataset. And the intervention intensity of each sub graph and sub series are always 4 or 5 time larger than intervention intensity of the original data.

Keywords: Intervention rule, Complex network, Time series, Decomposition, Multi-scale

1. Introduction

1.1. *Background*

In order to regulate financial market, reduce unemployment rate and control the spread of influenza, intervention is a widely used technique. However, before conducting intervention, the decision maker must make sure how much profit and loss the intervention may both engender. If profit is principal, the intervention should be executed. Otherwise the decision maker should try to carry out another intervention plain. Thus, intervention analysis has important economical value for the government.

In another aspect, mining multi-scale intervention rules has the following advantages.

Jiaoling Zheng,Changjie Tang, Shaojie Qiao, Ning Yang, Yue Wang,

(1) The data generated from complex system always contains information in different aspects. Thus if the data can be decomposed correctly, the interesting pattern hidden in the data can be mined out much more easily.

(2) As association rule can only reveal the co-appearance of two subjects, intervention rule may be more useful at revealing the casual and intervention relationship between two subjects.

1.2. *Motivation*

(a)The observation of multi-scale phenomenon

Figure 1 shows the breath rate and heart rate time series of a patient suffering from sleep apnea. The patient may occasionally suffocate while sleeping. Thus, the burst of breath rate will cause abnormal fluctuate of heart rate. However, if we just analyze the original data (Figure 1,a), no relationship between breath rate and heart rate is visibly explicit. If we retain only those points at odd timestamp intervals (Figure 1,b), it's easy to see that they exhibit similar patterns.
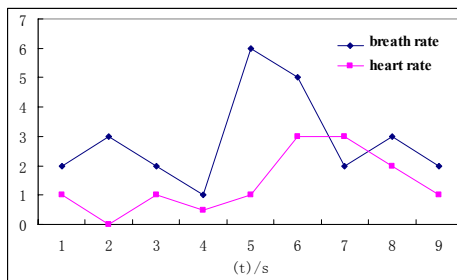


a. Original time series of breath rate and heart rate
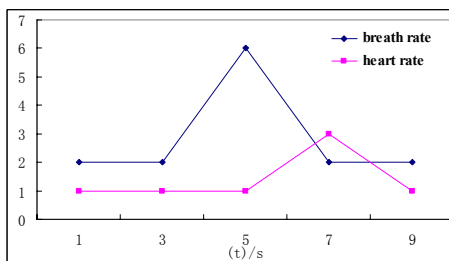
b.Decompose the original series into multi-scales to detect at what scale breath rate interferes with heart rate

Fig. 1. Multi-scale phenomenon

This observation shows that much more explicit relationship may exist at different decomposition scales of the original system.

(b)The quantification of intervention

From Figure 2(a), it is easy to see that the burst of breath rate will cause the burst of heart rate. But it is not quite convincing to say that breath rate interferes with heart rate, because not every burst of breath rate will cause the burst of heart rate. However, figure 2(b) shows that heart rate bursts each time after breath rate bursts. It is proper to say that breath rate interferes with heart rate.
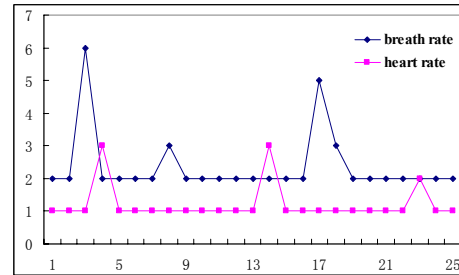
a. heart rate irregularly changes with breath rate
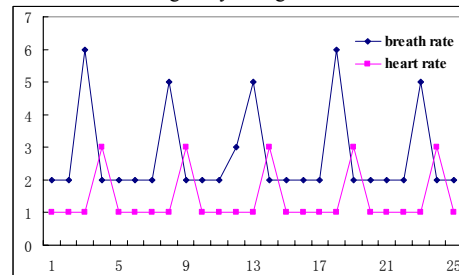
b. heart rate regularly changes with breath rate

Fig. 2. Intervention quantification

Zhang [1], et all, proposes the concept of naïve intervention rule. She uses the confidence index to quantify intervention intensity, i.e. intervention intensity = r_1 / r_2. r_1 is the burst times of heart rate and r_2 is the burst times of breath rate.

In this paper, we adopt correlation to measure intervention intensity, i.e. intervention intensity = correlation(X_1 , X_2). X_1 is the changing series of breath rate and X_2 is the changing series of heart rate.

Although there are a lot of numerical index to quantify intervention intensity, we consider correlation to be more appropriate. Since correlation is easy to calculate and requires strong liner relationship which may guarantee the intensity of the intervention rules.

(c)The direction of intervention

Association rule can reveal that the co-occurrence of breath rate and heart rate id frequent. But it can not tell which the cause is and which the consequence is. However, by using the intervention rule, it is easy to find that the intervention intensity from breath rate to

heart is 0.75 and intervention intensity from heart rate to breath is only 0.22. Thus it is rational to say that the burst of breath rate causes the burst of heart rate.

1.3. *Problem definition*

Definition 1 Intervention Rule Framework. Let X and Y be two elements in a system. Multi-scale intervention rule describe how X interferes with Y is a two tuple. Intervention(X→Y)=(Scale, Intensity). Scale denotes at what decomposition scale X and Y is investigated. Intensity = (correlation(ΔX_i , ΔY_i)), in which ΔX_i = (Δx_{i1}, Δx_{i2}, …, Δx_{in}) and ΔY_i = (Δy_{i1}, Δy_{i2}, …, Δy_{in}). Δx_{ij} ($1 \leq j \leq n$) denotes the jth change of X at decomposition scale i. Δy_{ij} ($1 \leq j \leq n$) denotes the jth change of Y according to Δx_{ij} at decomposition scale i.

This paper proposes the concept of intervention rule which tries to reveal the interventional relationship between elements in a system. Section 1 explains the idea of multi-scale intervention rules, and what are main differences between the intervention rules, association rules and causality analysis. Section 1 also introduces a general intervention rule framework Section 2 and section 3 transforms the framework into concrete intervention rules for complex network data and time series data and proposes two algorithms to mine the intervention rules from the two different systems. Section 4 shows the experimental results show that multi-scale intervention rules do exist in real dataset. Section 5 describes the related work and section 6 makes conclusion of this paper.

2. Mining Multi-scale Intervention Rules from Complex Network

Table 1. Denotation of complex network's intervention rule

Denotation	Meaning
X	X is a threshold, such that for any vertex v belonging to G, the degree of v is larger than X
Y	Y is the clustering coefficient of G, in which degree(v) is the degree of vertex v and mean is the mean value of all vertexes' degree values. $Y_{ij} = \sum |degree(v) - mean(Sub\_G_i)| (v \in G)$;
Scale	the sub graph of the original graph at the decomposition scale
Intensity	Intensity=Correlation(ΔX, ΔY) $\Delta X = (X_1, X_2, …, X_k)$ is the variation series of X; $\Delta Y = (Y_1, Y_2, …, Y_k)$ is the corresponding variation series of Y according to ΔX;

Definition 2 G = <V,E>: a citation network graph. Each paper v is a vertex in the graph G ($v \in V$). If one paper v is cited by another paper u, there will be an edge from u and v, denoted as $<u, v> \in E$.

According to definition 1, the two tuple Intervention(X→Y)=(Scale, Intensity) is the multi-scale intervention rule. The meaning of each denotation is described in table 1.

The goal of mining multi-scale intervention rules from citation network is to reveal how the change of X interferes with the change of Y in different sub graphs.

The following paragraphs will firstly introduce how to decompose a citation graph G into sub graphs, and then describe the algorithm of mining multi-scale intervention rule in each sub graph.

2.1. *Decomposition of the Original Graph*

The goal of decomposition is to divide the original graph G into several sub graphs such that each sub graph has much closer properties than the original graph. The decomposition algorithm is as follows.

Algorithm 1: Decompose(G)

Input: Original graph G
Output: Decomposed sub graphs of G {Sub_G_1, Sub_G_2,…, Sub_G_k,}

1. For each node v in G
2.   Vector(v) = <0,0,…,0>;//The length of Vector(v) //is |G|, where |G| is the number of nodes in G
3. End For
4. For each node v in G
5.   Vector(v) = Span(G,v);
6.   Num(v) = Count(Vector(v)); //return number of 1's in Vector(v)
7. End For
8. {Sub_G_1, Sub_G_2,…, Sub_G_k,}=Cluster(Num (v_1), Num(v_2),…, Num ($v_{|G|}$));
9. return {Sub_G_1, Sub_G_2,…, Sub_G_k,};

The following describes the meaning of algorithm1.

(1) Line 1~3 is to construct a zero vector Vector(v) for each node. The length of Vector (v) is the number of vertexes in graph G.

(2) Line 5 Span(G, v) is to traverse G in breath first manner with v as the starting node. If u (u≠v) is traversed, then the corresponding index of u in Vector(v) is set to 1.

(3) Line 8 is to conduct k-means upon Num (v_1), Num(v_2),…, Num ($v_{|G|}$). The clustering results will decompose G into several sub graphs. We also try other types of cluster methods. However, the empirical experimental results of these methods are

Jiaoling Zheng,Changjie Tang, Shaojie Qiao, Ning Yang, Yue Wang,

similar with k-means. So, we choose k-means as it is quite simple and well known. To determine the number of clusters by k-means, we assign k the value from 1 to K and choose the by which the algorithm can get minimum error rate.
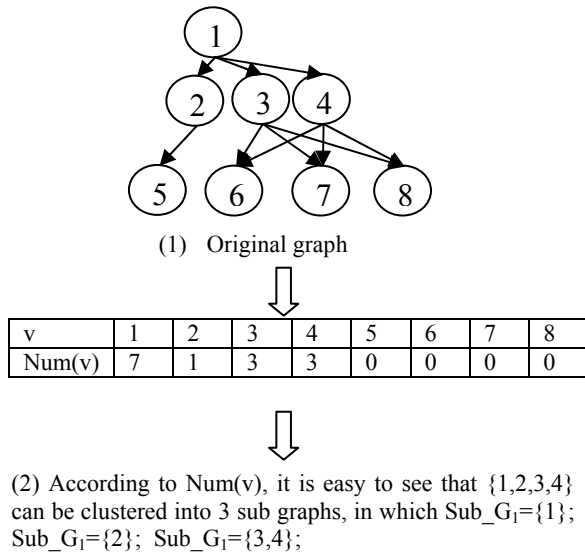

(1) Original graph

v	1	2	3	4	5	6	7	8
Num(v)	7	1	3	3	0	0	0	0

(2) According to Num(v), it is easy to see that {1,2,3,4} can be clustered into 3 sub graphs, in which Sub_G_1={1}; Sub_G_1={2}; Sub_G_1={3,4};

Fig. 3. Illustration of working process of algorithm 1

Example 1: Figure 3 shows the distribution of Num (v_1), Num(v_2),…, Num ($v_{|G|}$) of a citation network in the filed of HEP(high energy physics). The value of Num(v) of each node v is sorted in descending order. It is easy to see that the vertexes naturally form three clusters.

Hypothesis 1: Let {C_1, C_2,…, C_k} be the k clusters when the input of the k-means algorithm is the N×N sparse matrix {Vector(v_1), Vector(v_2),…, Vector($v_{|G|}$)}. For any C_i, C_j (i≠j), $|C_i|>>|C_j|$, or $|C_i|<<|C_j|$. |C| is the number of nodes in cluster C.
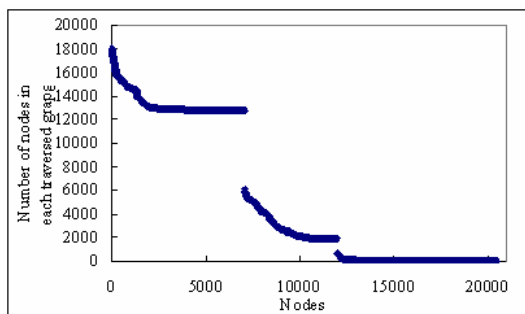

Fig. 4. Cluster of nodes in citation network

Explanation: According to the power law property of complex network, if we sort $|C_1|$, $|C_2|$,…, $|C_k|$ in descending order, the value will decrease exponentially. Thus $|C_i|>>|C_j|$, or $|C_i|<<|C_j|$. Example 2 shows such a phenomenon in the academic citation network.

Example 2: Figure 4 shows the distribution of nodes' degree in a citation network. It's easy to see that there are naturally three clusters.

Proposition 1: Let Result_1 be the clustering result with {Vector(v_1), Vector(v_2),…, Vector($v_{|G|}$)} as the input of k-means. Let Result_2 be the clustering result with {Num(v_1), Num(v_2),…, Num($v_{|G|}$)} as the input of k-means. Assume any node v∈Sub_G_i in Result_1 and v∈Sub_G_j in Result_2. Then, Sub_G_i≈Sub_G_j.

Proof: If u, v are in the same sub graph Sub_G_i in Result_1, then u, v must be in the same sub graph say Sub_G_j in Result_2. Because if u, v are in different sub graphs in Result_2, according to example 1, Num(u) will be much larger or smaller than Num(v). For example, if Num(u) is 1000, then Num(v) is 10000 or 100. Thus, Vector(u) and Vector(v) can not be in the same cluster i.e Sub_G_i in Result_1.

Equally, it is easy to prove that if u,v are in different sub graphs in Result_1, then u,v are likely to be in different sub graphs in Result_2. Thus, proposition 2 is proved.

Example 3: According to figure 4, it is easy to see that there are naturally three clusters {C_1, C_2, C_3}, in which $|C_1| \approx 15000$, $|C_2| \approx 4000$, $|C_1| \approx 10$. So if node u and v are clustered into C_1 when the input of k-means algorithm is {Vector(v_1), Vector(v_2),…, Vector($v_{|G|}$)}, it is not probable that u and v be clustered into different clusters when the input of k-means algorithm is { Num(v_1), Num(v_2),…, Num($v_{|G|}$)}.

Proposition 2: Let N = |G| be the number of nodes in graph G. Let Memory(Vector) and CPU(Vector) be the consumption of computer memory and the computing complexity when the input of the k-means clustering algorithm is {Vector(v_1), Vector(v_2),…, Vector(v_N)}. Let Memory(Num) and CPU(Num) be the consumption of computer memory and the calculating complexity when the input of the k-means clustering algorithm is { Num(v_1), Num(v_2),…, Num(v_N)}. We have Memory(Vector) = Memory(Num)^2 and CPU(Vector) = CPU(Num).

Proof: Memory(Vector) is an N×N matrix. Memory(Num) is an N×1 vector. So, it is easy to see

that Memory(Vector) = Memory(Num)^2. Num(i) = Vector(i,1) + Vector(i,2) +...+ Vector(i,N). According to proposition 1, it is easy to prove that the cluster results of the two different inputs are the same. Because the computing complexity of k-means algorithm is O(nkt) in which n is the number of input points, k is the number of clusters and t is the iterating time. Thus, CPU(Vector) = CPU(Num).

Example 4: According to figure 3, {Vector(v_1), Vector(v_2),…, Vector(v_N)} is an 7×7 matrix and { Num(v_1), Num(v_2),…, Num(v_N)} is an 7×1 vector. So, Memory(Vector) = Memory(Num)^2.And the two types of input both generate the same clustering result, i.e. Sub_G_1={1};   Sub_G_1={2};   Sub_G_1={3,4}. So, CPU(Vector) = CPU(Num).

2.2. *Mining Multi-scale Intervention Rules from Complex Network (MMIRCN)*

Section 2.1 has already described that the goal of intervention rule is to investigate how important nodes interfere with the clustering coefficient of a complex network. The intervention rule mining algorithm only returns those rules with intensity larger than a specified threshold, for example 0.5.

Algorithm 2(MMIRCN): **M**ining **M**ulti-scale **I**ntervention **R**ule from **C**omplex **N**etwork

Input: Original graph G, α // intensity threshold,
      Step_value// X's step value of variation ,
      k// number of sub graphs
      n//length of X and Y's series
Output: intervention rule set
1.   Rule_set=null;
2.   {Sub_G_1, Sub_G_2,…, Sub_G_k,}←Decompose(G);
3.   For i = 1 To k
4.   For j = 1 To n
5. X_{ij} = j*step_value;
6. Y_{ij}=∑|degree(v)-mean(Sub_G_i)|(v∈Sub_G_i);
7.   End For
8. ΔX_i=(X_{i1},X_{i2},…, X_{ik});
9. ΔY_i=(Y_{i1},Y_{i2},…, Y_{ik});
10.   Intensity = Correlation(ΔX_i, ΔY_i);
11.   Rule_set←<Sub_G_i, Intensity>;
12. End For
13. Return those rules in Rule_Set with intensity larger than α;

Line 2 is to decompose the original graph by algorithm 1. Line 3~ 12 is to generate intervention rules in each sub graph. Line 4~7 is to generate the varying sequence of <ΔX_i, ΔY_i >. Line 8~11 is to generate the intervention rule of each sub graph.

3. Mining Multi-scale Intervention Rules from Time Series

Definition 3 Time series : A *time series X* is a sequence of pairs (*timestamp; value*).The data values are ordered in timestamp ascending order. Let *s*(i) be the value of time series *x* at timestamp *i*, and x[i, j]=x(i)x(i+1)…x(j) be the subsequence of s at timestamp interval [i, j].

According to definition 1, the two tuple Intervention(X→Y)=(Scale, Intensity) is the multi-scale intervention rule. The meaning of each denotation is described in table 2.

Table 2. Denotation of time series' intervention rule

Denotation	Meaning
X	X is a time series in a complex system.
Y	Y is a time series in a complex system.
Scale	Scale denotes the sub series of the original series. Here, we adopt wavelet to decompose the original series.
Intensity	Intensity=Correlation(X, Y)

The goal of mining multi-scale intervention rules from time series is straight forward. Assume X and Y are two time series. Intervention(X→Y) is to measure how the change of X interferes with the change of Y. However, because the original series may be composed of multi-scale frequencies, it will be more fruitful to investigate the intervention at different decomposition scales.

The following paragraphs will firstly introduce how to decompose a time series into sub series, and then describe the algorithm of mining multi-scale intervention rule in each sub series.

3.1. *Decomposition of the Original Time Series*

This paper adopts wavelet transform to conduct decomposition on time series data. The decomposition algorithm is as follows.

Algorithm 3: Decompose(X)

Input: Original time series X=x(1)x(2)…x(L),
      //Assume the length of S is L
      k //total decomposition scales
Output: Decomposed sub series of X {Sub_X_1, Sub_X_2,…, Sub_X_k,}
1.   For i = 1 To k
2.   Sub_X_i = Wavelet_Decompose(X,i);
3.   End For
4.   Return {Sub_X_1, Sub_X_2,…, Sub_X_k,};

Wavelet_Decompose(X,i) is to decompose the original time series X into k different sub series. Sub_X_i is the sub series at decomposition scale i(1≤i≤k). One point need to empathies is the choice of mother wavelet.

Jiaoling Zheng,Changjie Tang, Shaojie Qiao, Ning Yang, Yue Wang,

Because different mother wavelets have different wavelet coefficients, not all kinds of wavelet transform are suitable for intervention rules mining. In Section 5.2, in order to calculate the directional correlation coefficients between two time series, we need to retain inflexion points of X in Sub_X$_i$. Only even symmetrical wavelet can satisfy this requirement. The odd symmetrical wavelet will transform the inflexion points of X into zero in Sub_X$_i$. So we choose discrete meyer wavelet [23] as mother wavelet. It is even symmetrical and orthogonal. The details of wavelet decomposition will be omitted here.

3.2. *Mining Multi-scale Intervention Rules from Time Series (MMIRTS)*

The goal of intervention rule by definition 1 is to investigate how the change of Sub_X$_i$ interferes with the change of Sub_Y$_i$ at scale i. The intervention rule mining algorithm only returns those rules with intensity larger than a specified threshold, for example 0.5.

Algorithm 4(MMIRTS): **M**ining **M**ulti-scale **I**ntervention **R**ule from **T**ime **S**eries

Input: time series X, Y
      k //total decomposition scales
      α // intensity threshold
Output: intervention rule set

1. Rule_set=null;
2. {Sub_X$_1$, …,Sub_X$_k$,}← Decompose(X);
3. {Sub_Y$_1$,…, Sub_Y$_k$,}← Decompose(Y);
4. For i = 1 to k
5.   Intensity = Correlation(Sub_X$_i$, Sub_Y$_i$);
6.   Rule_set←(<Sub_X$_i$, Sub_Y$_i$ >, Intensity);
7. End For
8. Return those rules in Rule_Set with intensity larger than α;

3.3. *Traditional Correlation vs Directional Correlation*

Intervention naturally has direction, i.e. X→Y or Y→X. But the traditional correlation can not represent the direction. So, this paper proposed two types of intervention intensity r$_1$ and r$_2$.

$$r_1 = TraCorr(X,Y) \qquad (1)$$

$$r_2 = DirCorr(X,Y) \qquad (2)$$

TraCorr is the traditional correlation, while DirCorr is directional correlation introduced by this paper. The following paragraphs will concentrate on explaining DirCorr.

Definition 4 Let X=X[1, t] be a time series. Then *a)* x(i) (1<i<t) is called *burst point* of X if |x(i)|>|x(i-1)| and |x(i)|>|x(i+1)|.

Definition 5 Let X[1, t] and Y=[1, t] be two time series with equal length t. The *Directional Correlation* between X and Y *(DirCorr(X, Y))* is defined as:

$$DirCorr(X,Y) = \sum_{k=1}^{m} w_{i,j} * Corr(X(i_k, j_k), Y(i_k, j_k))(w_{i,j} = \frac{j-i}{t}) \quad (3)$$

In Equation (3), X(i$_k$, j$_k$) needs to satisfy the following three conditions.(1) For any k, X(i$_k$), X(j$_k$) are burst point of X; (2) For any i$_k$≤p≤j$_k$, X(p) is not a burst point.(3)X(i$_1$, j$_1$) ∪ X(i$_2$, j$_2$) ∪ ·· ∪ X(i$_k$, j$_k$)=X(1,t);

Proposition 4. Let X=X[1,t]=(X(i$_1$,j$_1$),…, X(i$_m$,j$_m$)), Y=Y(1,t) =(Y(i$_1$,j$_1$),…, Y(i$_m$,j$_m$)), be two time series with equal length t.
Corr(X,Y)=r. And for any k, 1≤k≤m, X(i$_k$,j$_k$) is an monotonic subsequence of X(1,t) and For any k, 1≤k≤m,
(1) if Corr(X(i$_k$,j$_k$), Y(i$_k$,j$_k$))>r, then DirCorr(X,Y)>r;
(2) if Corr(X(i$_k$,j$_k$), Y(i$_k$,j$_k$))<r, then DirCorr(X,Y)<r;
Proof: Consider (1). By Equation 2, if Corr(X(i$_k$,j$_k$), Y(i$_k$,j$_k$))>r, we have DirCorr(X,Y)> w$_{i1,j1}$*r+…+ w$_{im,jm}$*r =r. Thus, (1) is proved. Similarly, it's easy to prove (2). According to proposition 1, The correlation of temporally adjacent points will not be weakened by other points. Moreover, because those temporally adjacent points are selected according to the burst points of X, thus, DirCorr(X,Y) can quantify the intervention from X to Y.

DirCorr can reveal the intervention direction in a way that DirCorr (X,Y) and DirCorr (Y,X) may be quite different. Example 5 shows the results.
Example 5: Let X[1, 8] and Y[1, 8] be sub series of patient's breath rate and heart rate in real dataset from time stamp 1 to time stamp 8.
X[1,8] =[-1.98, -2.41, -1.41, 0.42, 2.00, 2.41, 1.42, -0.41]
Y[1,8] =[-0.52, -2.20, -2.59, -1.45, 0.53, 2.21, 2.59, 1.46]
(1)DirCorr(X,Y)=2/8*Corr(X(1,2),Y(1,2))+4/8*Corr(X(3,6),Y(3,6))+2/8*Corr(X(7,8), Y(7,8))=0.97
(2)DirCorr(Y,X)=3/8*Corr(Y(1,3), X(1,3))+4/8*Corr(Y(4,7), X(4,7)) = 0.27
(3)Corr(X, Y) = 0.72
The results of directional correlation show that the intervention from X to Y is much stronger than the inverse direction. However, the traditional correlation is unable to reveal this relationship.

4. Experiments and Performance

The following paragraphs describe the experiment results on real datasets. All experiments are conducted in matlab6.5, Pentium 2.2G, 2G memory.

4.1. *Intervention rules of complex network*

The dataset of complex network is obtained from High Energy Physics (HEP)[2] data set. It contains the citation network of published papers in the filed of high energy physics from. It contains totally 20469nodes and 352785 edges.

From Figure 5, it is easy to see that the original citation network G can be decomposed into three sub graphs. The number of nodes in each sub graph is show in Table 3.
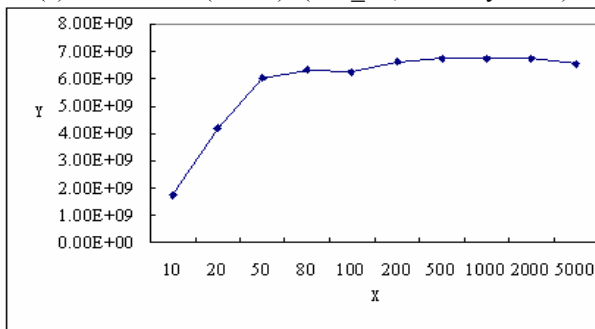
Table 3. Number of nodes in each sub graph

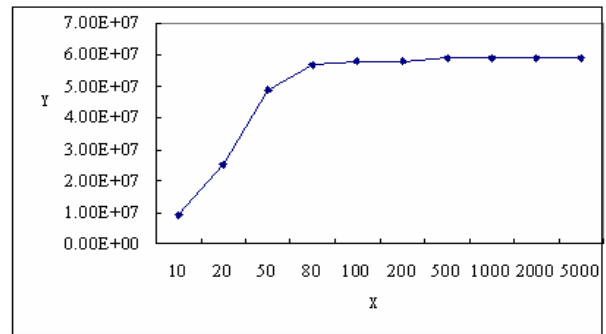| Sub graph of G	Number of nodes
Sub_G_1	7078
Sub_G_2	4898
Sub_G_3	8493

According to definition 1, the intervention rule of complex network is a two tuple Intervention(X→Y)= (Scale, Intensity).X and Y is defined in table 2.
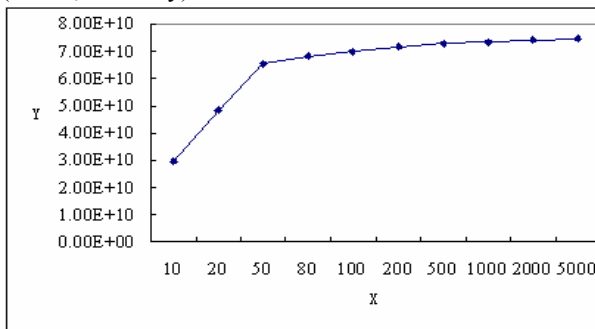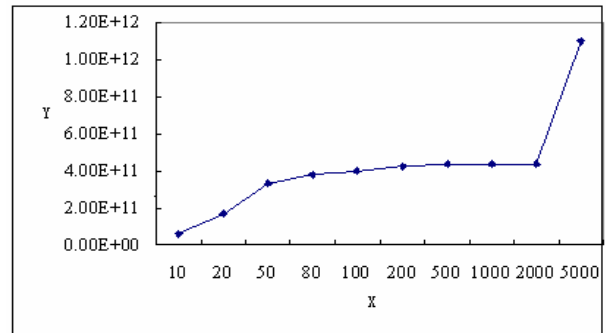

(a) Intervention(X→Y)=(Sub_G_1, Intensity=0.74)

(b) Intervention(X→Y)=(Sub_G_2, Intensity=0.68)

(c) Intervention(X→Y)=(Sub_G_3, Intensity=0.70)


(d) Intervention(X→Y)=(G, Intensity=0.85)

Fig. 5 Intervention rules in different decomposed sub graphs and the original graph

Figure 5 (a),(b),(c),(d) describes the change of Y according to the change of X in three sub graphs and in original graph. The corresponding intervention rule is shown under the figure. Because in each graph, Y and X exhibit logarithmic relationship, so Intensity = Correlation (Log(Y), Log(X)).

From the intervention rules, it is obvious that the X and Y value exhibit strong logarithmic relationship. Also, it is easy to see that both the original graph and each sub graph have strong intervention intensities. However, the cluster coefficient of each sub graph is at different exponential scale, i.e. 10^{10}, 10^9, 10^7. The intervention rule of the original graph can not reveal such subtle difference. Further more, the results accord with the theory of scale free network and show that those important nodes strongly interfere with the characteristics of complex network.

### 4.2. *Intervention rules of time series*

The dataset is obtained from Santa Fe time series [3] data set B1. It contains the breath rate and instantaneous heart rate of a sleeping human suffering from sleep apnea. We take the former 10000 records and normalize both series to zero mean and unit variance.
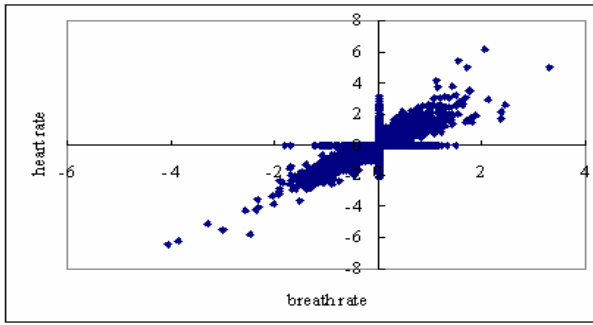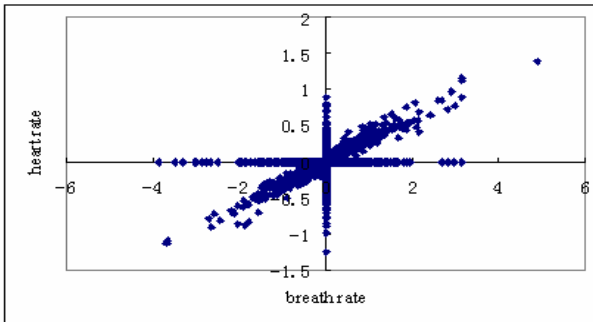
(a) Intervention(hr→br)=(sacle=1, r1=0.79)
Intervention(hr→br)=(sacle=1, r2=0.12)
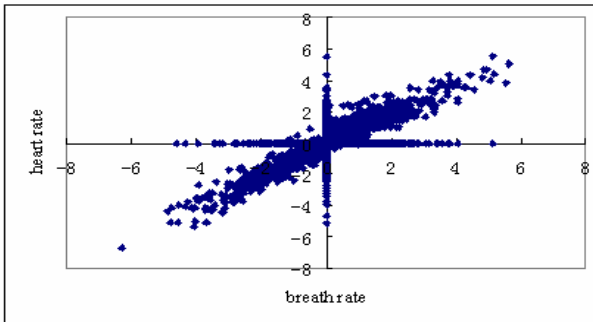
(b) Intervention(hr→br)=(sacle=2, r1=0.74)
Intervention(hr→br)=(sacle=2, r2=0.29)
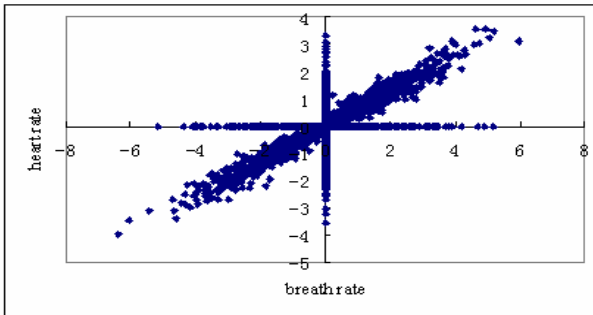
(c) Intervention(hr→br)=(sacle=3, r1=0.71)
Intervention(hr→br)=(sacle=3, r2=0.55)

(d) Intervention(hr→br)=(sacle=4, r1=0.56)
Intervention(hr→br)=(sacle=4, r2=0.16)

Fig. 6. Intervention rules from heart rate to breath rate

(a) Intervention(br→hr)=(sacle=1, r1=0.81)
Intervention(br→hr)=(sacle=1, r2= -0.20)

(b) Intervention(br→hr)=(sacle=2, r1= -0.77)
Intervention(br→hr)=(sacle=2, r2= -0.23)
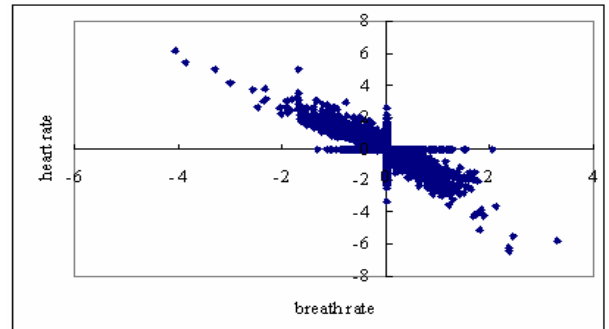

(c) Intervention(br→hr)=(sacle=3, r1= -0.69)
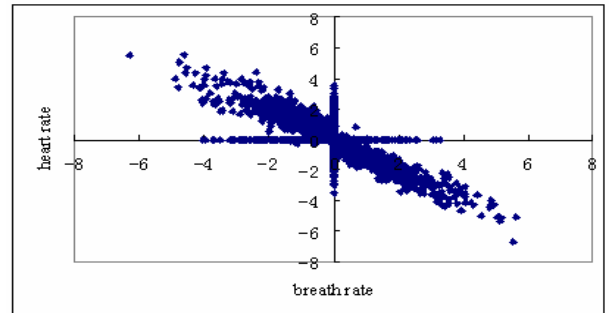Intervention(br→hr)=(sacle=3, r2= -0.47)

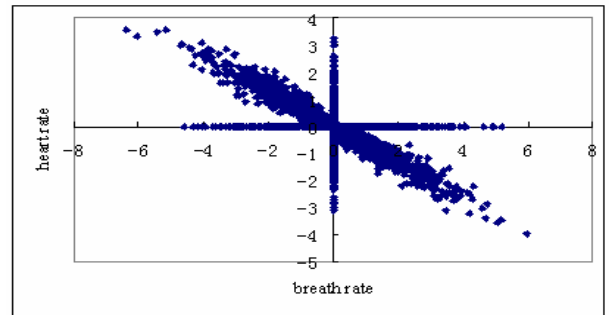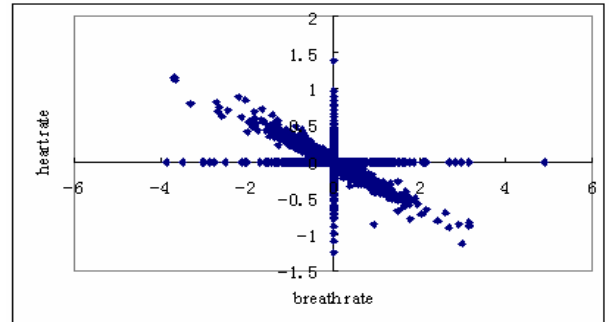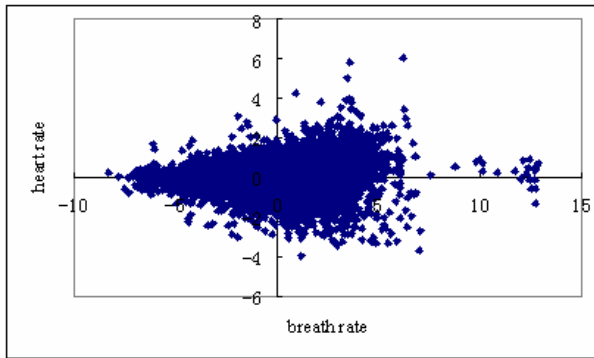(d) Intervention(br→hr)=(sacle=4, r1= -0.53)
Intervention(br→hr)=(sacle=4, r2= -0.14)

Fig. 7. Intervention rules from breath rate to heart rate

Intervention(br→hr)=(sacle= original, r1=0.07)
Intervention(br→hr)=(sacle= original, r2= 0.02)

Fig. 8. Intervention rules from breath rate to heart rate of the original time series

The original time series of the patient's breath rate and heart rate are decomposed into 4 different frequency scales. According to definition 1, the intervention rule of time series is a two tuple Intervention(X → Y)= (Scale, Intensity).

Figure 6 (a)~(d) show the change of breath rate with the change of heart rate at 4 different decomposition scales and the intervention rules from heart rate to breath rate. Figure 7 (a)~(d) show the change of heart rate with the change of breath rate at 4 different decomposition scales and the intervention rules from breath rate to heart rate. Figure 8 shows the intervention rules discovered in the original time series. We use hr as the abbreviated form of heart rate and br as the abbreviated form of breath rate. The corresponding intervention rule is shown under the figure and Intensity = Correlation (br,hr).

For each sub series, there are two types of intensities. r1 is the intensity using traditional correlation and r2 is the intensity using directional correlation. According to the intervention rules from heart rate to breath rate (Figure 6), although r1 in each decomposition scale are quite high, r2 is larger than the threshold 0.5 only in the 3rd scale.

According to the intervention rules from breath rate to heart rate (Figure 7), although r1 in each decomposition scale are quite high, r2 never go beyond the threshold 0.5 in all scales.

Thus, it is rational to conclude that the intervention direction is mainly from heart to breath. The results accord well with the experimental results in research [4] And also, the intensity in the original time series for both r1 and r2 are quite weak. Thus, it will be more fruitful to mine intervention rule at different scale.

4.3. *Performance of Algorithms*

(a) (MMIRCN): Mining Multi-scale Intervention Rule from Complex Network
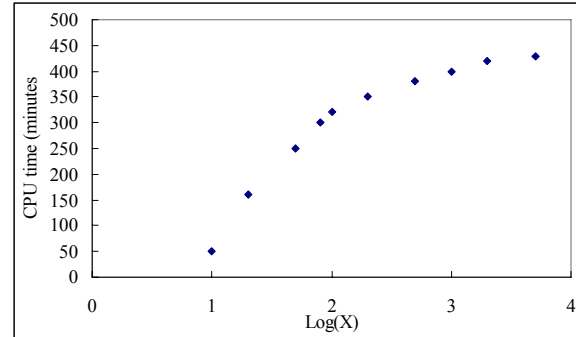
Fig 9. Performance of algorithms for MMIRCN

Figure 9 depicts the increase of computing time with the increase of X (X is defined in table 2). In algorithm 2, line 2 consumes the majority of the CPU time. It will traverse graph G' |Sub_G| times (|Sub_G| is the number of nodes in Sub_G). G' is the graph that don not contain those nodes with citation number larger than X. Thus if X is small, a lot of nodes will be remove from G to form G'. The computing time will be decreased.

(b) (MMIRTS): Mining Multi-scale Intervention Rule from Time Series
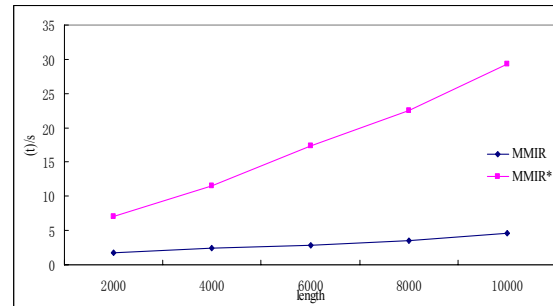
Fig 10. Performance of algorithms for MMIRTS

Figure 10 depicts the increase of computing time with the increase of time series' length. MMIR denotes the MMIRTS algorithm using traditional correlation. MMIR* denotes the MMIRTS algorithm using directional correlation.

Because calculating both the directional correlation only need to scan the whole time series constant times, so the complexity of MMIRTS using traditional correlation and directional correlation are both O(N). N is the length of time series.

Jiaoling Zheng,Changjie Tang, Shaojie Qiao, Ning Yang, Yue Wang,

5. Related Work

Intervention analysis is the common method to reveal relationships between objects in human as well as biological society. Data mining research community is just starting to pay attention to intervention analysis.

Tang et al. [5] irstly introduced the concept and the tasks of intervention rules mining, including intervention correlation, intervention type and intervention algebra. Zhang et al [1] proposed the naive intervention rules mining method based on association rules. Yang [6] proposed the algorithm MIPES to mine interventions in parallel event sequences by calculating the transfer entropy between time series. Guyon [7] employed intervention to improve Bayesian causal network of human T cells. His main idea is interfering with the active or non-active status of proteins to measure their effect on the whole protein network.

Because this paper focuses on mining intervention rules from complex network data and time series data, the following will briefly describe the relative work in the two fields.

Literatures [8~11] researched the hidden relationship between temporal sequences via pattern matching, such as biological, medical and economical time series. Several other works [12~15] are devoted to measure the strength and direction of information flow between simultaneously observed time series.

In another aspect, multiple resolution analysis (MRA) [16] and wavelet transform [17] are widely used in time series analysis, such as dimension reduction or signal compression. However, as correlation can not indicate the direction of intervention, this study introduces directional correlation.

Multi-scale networks are a common in many social and biological complex networks. The early research results of [18] shows that the river networks is hierarchical and even exhibit fractural property. Cardiovascular network, the roots, leaf veins and branches of trees all exhibit fractural properties [19]. P.S. Dodds investigated the spreading of information over hierarchical networks [20]. He concluded that multi-scale networks are ultra-robust and scalable [21]. Slater's research shows that the backbones many practical networks, such as traffic network, the network of provinces of a country, exhibit multi-scale properties [22]. Research [23] investigates the scaling and multi-scale behavior of traffic network and uses wavelet to predict the incoming traffic. Chun-Biu Li presents a

scheme to extract a multi-scale state space network from a single-molecule time series to lift degeneracy in molecular study as much as possible [24]. Shino proposes new type of point-pattern analytical method to identify point agglomerations across multi-scale network-based clumps among distributed points along a network [25]. Qiao et al. discovered frequent trajectory patterns from multi-scale trajectory networks [26, 27].

However, although multi-scale network is wildly studied, the mining of intervention rules from those networks is leaving unconcerned. As the properties of different complex network vary a lot, the decomposition methods are different. Because this study focuses on decomposing the citation network, the width first traversing algorithm is adopted for decomposition.

The organization of the remained sections is as follows. (1) Section 4 introduces how to mining multi-scale intervention rules from complex network. (2) Section 5 introduces how to mining multi-scale intervention rules from time series data. (3) Section 6 shows the multi-scale intervention rules discovered and the performances of the algorithms. (4) Section 7 describe conclusions and future work.

6. Conclusion and Future work

This study aims to mine quantitative intervention rules from complex network and time series data. The main contributions include: (a) defined new concepts of multi-scale intervention rules both for network and time series; (b) conducted decomposition to divide the original data into multiple scales; (c) proposed algorithms to mine intervention rules from those decomposed sub data. (d) conducted experiments to show that the proposed method successfully find intensive intervention rules.

The future work is to apply the idea of multi-scale analysis to more interesting biological data such as human brain wave, micro array data of DNA. In addition, we will employ other intelligent information processing techniques as presented in [28-30] to mine intervention rules from Complex Systems.

References

1. Y. Zhang, C. Tang, C. Li, J. Zhu, C. Zeng, L. Tang, X. Liu "Mining Naive Intervention Rules in Birth Defect Data".*Journal of Frontiers of Computer Science and Tehnology*. 1(2), 188-197,(2009)(in Chinese)
2. J. Leskovec, J. Kleinberg, C. Faloutsos. "Graphs over Time: Densification Laws, Shrinking Diameters and Possible Explanations". *ACM SIGKDD International*

Conference on Knowledge Discovery and Data Mining,177-187,2005.

3. D. R. Rigney, "Time Series Prediction: Forecasting the Future and Understanding the Past", *International Institutes of forcasters*,10(3),463-466,(1994)

4. Thomas Schreiber," Measuring information transfer" *Physics Review Letter*. 85(2),461–464, (2000)

5. C. Tang, Y. Zhang, L. Tang, C. Li, J. Peng. "A Survey on Mining Kinetic Intervention Rule from Sub-complex Systems" *Journal of Computer Application*, **28**(11), 2732-2736+2748 (2008) (in Chinese).

6. N. Yang, C. TANG, Y. Wang, "Mining Interventions from Parallel Event Sequences", *Proceedings of the Joint International Conferences on Asia-Pacific Web Conference and Web-Age Information Management*, 4, 297-307 (2009).

7. I. Guyon, C. Aliferis, G. Cooper, A. Elisseeff, J. Pellet, P. Spirtes, A. Statnikov "Design and Analysis of the Causation and Prediction Challenge", *Proceedings of World Congress on Computational Intelligence*, 3,1-33 (2008)

8. J. Roddick, K. Hornsby, "An Updated Bibliography of Temporal, Spatial and Spatio-Temporal Data Mining Research" *Proceedings of International Conference on Solid State Devices and Materials*, 2,147--163, (2007)

9. O. Antunes, "Temporal data mining: An overview". *In: Workshop on Temporal Data Mining, 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,5,1-13* (2001).

10. J. Han, H. Cheng, D. Xin, X. Yan, "Frequent Pattern Mining: Current Status and Future Directions". *Data Mining and Knowledge Discovery*, 15(1),55–86, (2007)

11. K. P. Unnikrishnan, R. Uthurusamy, "Temproal Data Mining" *Proceedings of 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* ,6(2), 205-217, (2001)

12. M.G. Rosenblum, A.S. Pikovsky, "Detecting direction of coupling in interacting oscillators", *Physics Review*, 64(4), 45-202,(2001)

13. J. Arnhold, P. Grassberger, K. Lehnertz, C.E. Elger, "A robust method for detecting interdependences application to intracranially recorded EEG", *Physica D, 2(*134), 419-430,(1999)

14. M.G. Rosenblum, A.S. Pikovsky, J. Kurths, "Phase synchronisation of chaotic attractors", *Physics Review Letter*. 76(11),1804-1807, (1996)

15. M. Le Van Quyen, C. Adam, M. Baulac, J. Martinerie, F.J. Varela, "Non-linear interdependences of EEG signals during intracranial ictal activities", *Brain Research*. 792(2),24-40,(1998)

16. S. Mallat, "A Theory for Multi-resolution Signal Decomposition: The Wavelet Representation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*,7 (11), 674-693,(1989)

17. S. Mallat, "Multi-resolution Representation and Wavelets," *PhD Thesis, Univ. of Pennsylvania, Philadelphia*, 1988

18. Ignacio Rodriguez Iturbe, "Fractal River Basins Chance and Self-Organization" Cambridge University Press, Cambridge UK, (1997)

19. M. Zamirf, "On Fractal Properties of Arterial Trees", *Journal of Theoretical Biology,*197(4), 517-526,(1999)

20. P.S. Dodds, D.H. Rothman, "Scaling, universality, and geomorphology", *Annual Review of Earth and Planetary Sciences*, 28(2), 571-610, (2000)

21. S. Dodds, J. Watts, F. Sabel, "Ultra-Robust and Scalable Networks Based on Hierarchies", *Computing Science and Statistics*,11(7), 35-49, (2003)

22. P. B. Slater, "A two-stage algorithm for extracting the multiscale backbone of complex weighted networks". *Proc Natl Acad Sci,*106(26),38-59,2009

23. F. Hong, Z. Wu, "Multiscale network traffic prediction model based on wavelet". *Chinese Journal of Computers*. 29(1),166-170,(2006)

24. C. Li, H. Yang, T. Komatsuzaki, "Complex Network of Protein Conformational Fluctuation Buried in Single Molecule Time Series", *Proceedings of the National Academy of Sciences*, 105(2),536-541,(2008)

25. S. Shiode, N. Shiode, "Detection of multi-scale clusters in network space", *International Journal of Geographical Information Science*, 23(1),75-92, 2009

26. S. Qiao, T. Li, J. Peng, J. Qiu, "Parallel sequential pattern mining of massive trajectory data," *International Journal of Computational Intelligence Systems*, **3**(3), 343-356 (2010).

27. S. Qiao, C. Tang, H. Jin, T. Long, S. Dai, Y. Ku, M. Chau, "PutMode: prediction of uncertain trajectories in moving objects databases," *Applied Intelligence*, **33**(3), 370-386 (2010).

28. S. Qiao, C. Tang, H. Jin, J. Peng, D. Davis, N. Han, "KISTCM: knowledge discovery system for traditional Chinese medicine," *Applied Intelligence*, **32**(3), 346-363 (2010).

29. S. Qiao, T. Li, M. Chau, J. Peng, Y. Zhu, H. Li, "WebScore: an effective page scoring approach for uncertain web social networks," *International Journal of Computational Intelligence Systems*. (unpublished)

30. S. Qiao, T. Li, H. Li, H. Chen, J. Peng, J. Qiu, "HCUBE: a hierarchical clustering algorithm using blockmodeling in web social networks," *Proceedings of the 9th International FLINS Conference on Foundations and Applications of Computational Intelligence*, **4**, 697-702 (2010).