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

Jiaoling Zheng¹, Changjie Tang², Shaojie Qiao³, Ning Yang², Yue Wang²

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

Received 2 September 2010; accepted 18 August 2011

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 times 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.

*This work was supported by the National Science Foundation of China under Grant Nos. 61100045, 61165013, 60773169; the Youth Foundation for Humanities and Social Sciences of Ministry of Education of China under Grant No. 10YJCZH117; the National Science Foundation for Post-doctoral Scientists of China under Grant No. 20090461346; the Fundamental Research Funds for the Central Universities under Grant No. SWJTU09CX035; the Scientific Research Foundation of CUIT under Grant No. KYTZ201110; the 11th Five Years Key Programs for Sci.&Tech. Development of China under Grant No. 2006BAI05A01.

† Corresponding author: sjqiao@home.swjtu.edu.cn

Published by Atlantis Press
Copyright: the authors

728
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.

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.

(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.

Zhang [1], et all, proposes the concept of naïve intervention rule. She uses the confidence index to quantify intervention intensity, i.e. intervention intensity = r1 / r2. r1 is the burst times of heart rate and r2 is the burst times of breath rate.

In this paper, we adopt correlation to measure intervention intensity, i.e. intervention intensity = correlation( X1 , X2 ). X1 is the changing series of breath rate and X2 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 linear 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( ΔX1, ΔY1, ), in which ΔXi = (Δx1i, Δx2i, ..., Δxni) and ΔYi = (Δy1i, Δy2i, ..., Δyni). Δxi (1≤i≤n) denotes the jth change of X at decomposition scale i. Δyi (1≤i≤n) denotes the jth change of Y according to Δxj 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

<table>
<thead>
<tr>
<th>Denotation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>X is a threshold, such that for any vertex v belonging to G, the degree of v is larger than X</td>
</tr>
<tr>
<td>Y</td>
<td>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. Yv=sum(degree(v)-mean(Sub_Gi))/(v∈G);</td>
</tr>
<tr>
<td>Scale</td>
<td>the sub graph of the original graph at the decomposition scale</td>
</tr>
<tr>
<td>Intensity</td>
<td>Intensity=Correlation(ΔX, ΔY)</td>
</tr>
<tr>
<td>ΔX</td>
<td>ΔX=(X1,X2, ..., Xk) is the variation series of X;</td>
</tr>
<tr>
<td>ΔY</td>
<td>ΔY=(Y1,Y2, ..., Yk) is the corresponding variation series of Y according to ΔX;</td>
</tr>
</tbody>
</table>

Definition 2 G = <V,E>: a citation network graph. Each paper v is a vertex in the graph G (v ∈ V). If one paper v is cited by another paper u, there will be an edge from u and v, denoted as <u, v> ∈ 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_G1, Sub_G2, ..., Sub_Gk,}

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));
7. //return number of 1’s in Vector(v)
8. {Sub_G1, Sub_G2,..., Sub_Gk,} = Cluster(Num (v1), Num(v2),... Num (v|G|));
9. return {Sub_G1, Sub_G2,..., Sub_Gk,};

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 (v1), Num(v2),... 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
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.

**Example 1:** Figure 3 shows the distribution of Num (v1), Num(v2),…, 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(v1), Vector(v2),…, 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.

**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(v1), Num(v2),…, 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| ≈ 15000, |C 2| ≈ 4000, |C 3| ≈ 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 |G|)}. 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 |G|) }. 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(v1), Vector(v2),..., Vector(vN)} is an 7×7 matrix and { Num(v1), Num(v2),..., Num(vN)} 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_G1={1}; Sub_G1={2}; Sub_G1={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): Mining Multi-scale Intervention Rule from Complex Network

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_G1, Sub_G2,..., Sub_Gk,}←Decompose(G);
3. For i = 1 To k
4. For j = 1 To n
5. Xij = j*step_value;
6. Yij = \sum(deg(v)-mean(Sub_Gi))(v ∈ Sub_Gi);
7. End For
8. ∆X_i=(X_1,X_2,...,X_k);
9. ∆Y_i=(Y_1,Y_2,...,Y_k);
10. Intensity = Correlation(∆X_i, ∆Y_i);
11. Rule_set←<Sub_Gi, 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 <∆X_i, ∆Y_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

<table>
<thead>
<tr>
<th>Denotation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>X is a time series in a complex system.</td>
</tr>
<tr>
<td>Y</td>
<td>Y is a time series in a complex system.</td>
</tr>
<tr>
<td>Scale</td>
<td>Scale denotes the sub series of the original series. Here, we adopt wavelet to decompose the original series.</td>
</tr>
<tr>
<td>Intensity</td>
<td>Intensity=Correlation(X, Y)</td>
</tr>
</tbody>
</table>

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_X1, Sub_X2,..., Sub_Xk,}

1. For i = 1 To k
2. Sub_Xi = Wavelet_Decompose(X,i);
3. End For
4. Return {Sub_X1, Sub_X2,..., Sub_Xk,};

Wavelet_Decompose(X,i) is to decompose the original time series X into k different sub series. Sub_Xi is the sub series at decomposition scale i(1≤i≤k). One point need to empathies is the choice of mother wavelet.
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 \( \text{Sub}_X \). Only even symmetrical wavelet can satisfy this requirement. The odd symmetrical wavelet will transform the inflexion points of \( X \) into zero in \( \text{Sub}_X \). 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 \( \text{Sub}_X \) interferes with the change of \( \text{Sub}_Y \) 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): Mining Multi-scale Intervention Rule from Time Series

**Input:** time series \( X, Y \)

\[ k \] / total decomposition scales

\[ \alpha \] // intensity threshold

**Output:** intervention rule set

1. \( \text{Rule}\_\text{set}=\emptyset \);
2. \( \{\text{Sub}_X_1, \ldots, \text{Sub}_X_k\} \leftarrow \text{Decompose}(X) \);
3. \( \{\text{Sub}_Y_1, \ldots, \text{Sub}_Y_k\} \leftarrow \text{Decompose}(Y) \);
4. For \( i = 1 \) to \( k \)
5. \hspace{1em} \text{Intensity} = \text{Correlation}(\text{Sub}_X_i, \text{Sub}_Y_i) ;
6. \hspace{1em} \text{ Rule } \_\text{set} \leftarrow (\text{Sub}_X_i, \text{Sub}_Y_i, \text{Intensity} ) ;
7. End For
8. Return those rules in \( \text{Rule}\_\text{set} \) with intensity larger than \( \alpha \).

#### 3.3. Traditional Correlation vs Directional Correlation

Intervention naturally has direction, i.e. \( X \rightarrow Y \) or \( Y \rightarrow X \). But the traditional correlation can not represent the direction. So, this paper proposed two types of intervention intensity \( r_1 \) and \( r_2 \).

\[ r_1 = \text{TraCorr}(X, Y) \]  
\[ r_2 = \text{DirCorr}(X, Y) \]  

\( \text{TraCorr} \) is the traditional correlation, while \( \text{DirCorr} \) is directional correlation introduced by this paper. The following paragraphs will concentrate on explaining DirCorr.

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

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

\[
\text{DirCorr}(X, Y) = \sum_{i=1}^{t} w_{ij} \cdot \text{Corr}(X(i, j), Y(i, j))(w_{ij} = \frac{j-1}{t})
\]  

In Equation (3), \( X(i, j) \) needs to satisfy the following three conditions. (1) For any \( k \), \( X(i, j) \) is burst point of \( X \); (2) For any \( i \leq p \leq j \), \( X(p) \) is not a burst point. (3) \( X(i, j) \) \( \cup X(i, j) \) \( \cup \cdots \cup X(i, j) = X(t) \).

**Proposition 4.** Let \( X = [1, t] = (X(i_1, j_1), \ldots, X(i_m, j_m)) \), \( Y = [1, t] = (Y(i_1, j_1), \ldots, Y(i_m, j_m)) \), be two time series with equal length \( t \).

\[ \text{Corr}(X, Y) = r \]  

And for any \( k \), \( 1 \leq k \leq m \), \( X(i_k, j_k) \) is an monotonically increasing subsequence of \( X(t) \) and For any \( k \), \( 1 \leq k \leq m \), (1) if \( \text{Corr}(X(i_k, j_k), Y(i_k, j_k)) > r \), then \( \text{DirCorr}(X, Y) > r \); (2) if \( \text{Corr}(X(i_k, j_k), Y(i_k, j_k)) < r \), then \( \text{DirCorr}(X, Y) < r \).

**Proof:** (Consider 1). By Equation 2, if \( \text{Corr}(X(i_k, j_k), Y(i_k, j_k)) > r \), we have \( \text{DirCorr}(X, Y) > w_{1,1} \cdot r + \cdots + w_{m, m} \cdot r \).

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, \( \text{DirCorr}(X, Y) \) can quantify the intervention from \( X \) to \( Y \).

\( \text{DirCorr} \) can reveal the intervention direction in a way that \( \text{DirCorr}(X, Y) \) and \( \text{DirCorr}(Y, X) \) 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) \( \text{DirCorr}(X, Y) = 2/8 \cdot \text{Corr}(X(1, 2), Y(1, 2)) + 4/8 \cdot \text{Corr}(X(3, 6), Y(3, 6)) + 2/8 \cdot \text{Corr}(X(7, 8), Y(7, 8)) = 0.97 \)

(2) \( \text{DirCorr}(Y, X) = 3/8 \cdot \text{Corr}(Y(1, 3), X(1, 3)) + 4/8 \cdot \text{Corr}(Y(4, 7), X(4, 7)) = 0.27 \)

(3) \( \text{Corr}(X, Y) = 0.72 \)

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) data set. It contains the citation network of published papers in the field of high energy physics from. It contains totally 20469 nodes 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 shown in Table 3.

<table>
<thead>
<tr>
<th>Sub graph of G</th>
<th>Number of nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub _G1</td>
<td>7078</td>
</tr>
<tr>
<td>Sub _G2</td>
<td>4898</td>
</tr>
<tr>
<td>Sub _G3</td>
<td>8493</td>
</tr>
</tbody>
</table>

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 _G1, Intensity=0.74)
(b) Intervention(X→Y)=(Sub _G2, Intensity=0.68)
(c) Intervention(X→Y)=(Sub _G3, Intensity=0.70)
(d) Intervention(X→Y)=(G, Intensity=0.85)

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 = \text{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 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.
Fig. 6. Intervention rules from heart rate to breath rate

(a) $\text{Intervention}(hr \rightarrow br) = (sacle=1, r_1=0.79)$
$\text{Intervention}(hr \rightarrow br) = (sacle=1, r_2=0.12)$

(b) $\text{Intervention}(hr \rightarrow br) = (sacle=2, r_1=0.74)$
$\text{Intervention}(hr \rightarrow br) = (sacle=2, r_2=0.29)$

(c) $\text{Intervention}(hr \rightarrow br) = (sacle=3, r_1=0.71)$
$\text{Intervention}(hr \rightarrow br) = (sacle=3, r_2=0.55)$

(d) $\text{Intervention}(hr \rightarrow br) = (sacle=4, r_1=0.56)$
$\text{Intervention}(hr \rightarrow br) = (sacle=4, r_2=0.16)$

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

(a) $\text{Intervention}(br \rightarrow hr) = (sacle=1, r_1=0.81)$
$\text{Intervention}(br \rightarrow hr) = (sacle=1, r_2=-0.20)$

(b) $\text{Intervention}(br \rightarrow hr) = (sacle=2, r_1=-0.77)$
$\text{Intervention}(br \rightarrow hr) = (sacle=2, r_2=-0.23)$

(c) $\text{Intervention}(br \rightarrow hr) = (sacle=3, r_1=-0.69)$
$\text{Intervention}(br \rightarrow hr) = (sacle=3, r_2=-0.47)$

(d) $\text{Intervention}(br \rightarrow hr) = (sacle=4, r_1=-0.53)$
$\text{Intervention}(br \rightarrow hr) = (sacle=4, r_2=-0.14)$
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

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

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.
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.


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 widely 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

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


7. 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)


