差异区域平衡法探索时间序列变化的因果关系

王开军* 曾元鹏 缪忠剑
(福建师范大学数学与信息学院 福州 350117)
(福建师范大学数字福建环境监测物联网实验室 福州 350117)

摘 要：针对探索时间序列之间随时间变化的因果关系问题，在每个窗口进行Granger因果检测的滑动时间窗口方法是求解该问题的常用方法，但其性能对窗宽敏感，不合适的窗宽很可能导致低性能。该文提出一种差异区域平衡方法，首先计算当前滑动窗口W内序列的波动程度Sw并作为波动界，计算窗口W的前向相邻区域U内序列的波动程度Su。然后，实施前向探索策略：若Su未超过Sw，则实施不同长度区域的平衡检测方案，即对窗口W、对窗口W与U的合并区域、对窗口W与后向相邻区域V的合并区域这3种不同长度的差异区域，分别进行时间序列之间因果关系的检测；若Su超过Sw，则实施上述平衡检测方案时，其中区域U和V的长度取相同值。最后，将窗口W的多次检测结果进行综合后输出。新方法将不同长度区域的结果进行综合，能够降低方法的性能对窗宽的敏感性，保障最终结果的准确性和稳定性。在1个模拟数据集和4个真实数据集上的实验结果显示，该文方法能有效地揭示出时间序列之间随时间变化的因果关系，在正确率高且性能稳定的综合性能上优于对比方法。

关键词：时间序列；变化的因果关系；Granger因果检测；差异区域平衡
中图分类号：TP391.4 **文献标识码**：A **文章编号**：1009-5896(2021)08-2414-07
DOI: 10.11999/JEIT200756

Different-region Balance Method for Exploring Varying Causal Relations Between Time Series

WANG Kaijun ZENG Yuanpeng MIAO Zhongjian
(*College of Mathematics and Informatics*, *Fujian Normal University*, *Fuzhou* 350117, *China*)
(*Digit Fujian Internet-of-Things Laboratory of Environmental Monitoring*, *Fuzhou* 350117, *China*)

Abstract: For discovering time-varying causal relations between time series, a common method is the sliding-window method with Granger causal tests on every window. However, the method performance is sensitive to window sizes, and an unsuitable size probably leads to poor performance. The different-region balance method is proposed. The variation degree of time series in current sliding window W (called variation bound Sw) is first computed, and the degree Su in front neighbor region U of W is computed. Then a forward exploring strategy is adopted: when Su≤Sw, a different-length-region balance test measure is carried out, i.e., causal-relation tests respectively in window W, combined region W and U, and combined window W and back neighbor region V of W; when Su>Sw, it uses the above-mentioned measure where region V has the same length as region U; Finally, in each region, all the test results are synthesized to give a final result. The new method combines the results from different-length regions to reduce its sensitivity to window sizes, and guarantees the accuracy and stability of final results. The experiments on one simulated data set and four real data sets show that, the new method can discover time-varying causal relations between time series effectively, and outperforms the compared methods on the balance performance of high accuracy and stability.

Key words: Time series; Time-varying causal relations; Granger causal test; Different-region balance

收稿日期：2020-08-26；改回日期：2021-01-01；网络出版：2021-01-07
*通信作者： 王开军 wkjwang@qq.com
基金项目：国家自然科学基金(61672157)，福建省自然科学基金 (2018J01778)
Foundation Items: The National Natural Science Foundation of China (61672157), The Natural Science Foundation of Fujian Province (2018J01778)

1　引言

发现和识别事物之间的因果关系是许多研究领域的重要主题[1,2]，而探索时间序列之间的因果关系对于研究动态发展的系统具有重要的理论和实际意义[3]。已有一些挖掘时间序列间因果关系的方法，例如状态空间方法[4]通过建立状态空间模型描述具有时间序列的动态系统，然后根据测量方程来判别因果关系，文献[5]在回归模型中并入时滞序列数据，然后用条件概率和期望传播方法求解模型参数来获得因果关系。针对2元时间序列的因果关系分析问题，经典的Granger因果分析方法[6]的解决思路是：当某个序列变量加入Granger因果模型有助于获得更好的预测结果时，可认为模型中的变量之间具有统计意义上的因果关系。基于时间序列的Granger因果分析具有好的可解释性，在许多领域得到了应用，例如Granger因果分析方法用于分析阻塞性睡眠呼吸暂停患者的多频道睡眠时间序列来检测脑电波和心率变异性之间的关联性[7]，用于探究我国经济发展和社会发展的因果关系[8]。针对多元时间序列，文献[9]结合蒙特卡罗方法提出了基于多变量的Granger因果分析方法，文献[10]提出了基于Hilbert-Schmidt独立准则的Lasso Granger因果分析方法来分析多元时间序列之间的非线性因果关系，更多关于多变量的Granger因果分析方法参见文献[3]。

对于随时间而发展变化的系统，其时间序列(序列变量)之间的因果关系可能随时间而发生变化。探寻时间序列中发展变化的因果关系是尚待解决的困难问题，有关的研究文献较少。不失一般性，这里讨论两条时间序列之间的线性因果关系随着时间改变的情况，侧重于因果关系随时间的变化问题，例如它们之间的因果关系在某个时间区间存在，而在下一个时间区间则变化为不存在。常规的滑动窗口方法[11,12]采用滑动时间窗口搜索时间序列，并用Granger因果方法检测当前时间窗口内时间序列之间的因果关系；但是，窗宽和步长的不同会导致结果和性能的差异。滑动窗口方法若穷举所有可能的时间区间(遍历方式)进行因果关系探寻时，对于长度为T、时滞为L的两条时间序列，其时间复杂度高达$O(T^3L^2)$[13]。F界检测法[13]采用遍历方式对包含时刻t的所有不同长度的时间区间进行因果关系检测，其中对已检测过的区间的相邻区域使用F检验的上下界判别因果关系；对每个t计算包含t的因果区间数量占全部区间数量的比值作为因果分，依据因果分可判别不同时间区间上序列间因果关系是否存在。文献[14]采用转折点检测算法找出时间序列上的转折点(作为分割点)，然后在分割区域上进行线性Granger因果检测，获得随时间演化的关系网络(简称转折点检测法)；该方法认为由转折点分割出的不同时间区域在产生时间序列的机制上对应着因果关系结构的改变。但是，复杂多变的时间序列会干扰这种转折点的正确识别，进而导致该方法性能的下降。可以看出，常规滑动窗方法的性能对不同窗宽和步长比较敏感而带来性能不稳定；F界检测法遍历不同长度的时间区间的方式偏重全局信息，局部信息倾向于被弱化，则在信号弱的局部区域上容易出现误判导致性能降低；转折点检测法的性能很容易被时间序列数据的波动干扰而降低；因而上述方法在探寻时间序列间发展变化的因果关系上的性能有待进一步改进。

针对滑动窗方法的不足，本文提出一种差异区域平衡方法，设计了前向探索策略、不同长度区域的平衡检测方案和多数票赢的结果综合等措施，探索时间序列之间的因果关系随时间发展而发生变化的情况，期待在合理的窗宽步长时新方法具有更好的性能且稳定。

2　时间序列因果发现的相关方法

常规滑动窗方法探索时间序列随时间变化的关系的基本框架为：宽wide的滑动窗Wt按步长step沿时间t方向移动，同时在当前窗口Wt内对时间序列间关系进行探查，并采用Granger因果检测方法探查窗口内时间序列之间是否存在因果关系，记录探查结果。

Granger因果分析的基本思想是：如果时间序列X引起了时间序列Y的变化，相较于仅使用Y的历史数据预测Y的未来值(常用自回归模型描述，称约减模型Mr)，使用X和Y的历史数据对Y的未来值进行预测(常用回归模型描述，称完整模型Mf)，将提高其预测准确性。下文简介实现上述思想的Granger因果检测方法。

给定长度为 T 的时间序列$X=\{x_1,x_2,\cdots,x_T\}$和$Y=\{y_1,y_2,\cdots,y_T\}$，拟合时间序列X和Y的两种回归模型[13]为(时滞参数L、回归参数a和b)

$$\left.\begin{aligned}&\text{Mf}:\quad \hat{y}_t^{\text{f}}\sim\sum_{l=1}^{L}a_l\cdot y_{t-l}+\sum_{l=1}^{L}b_l\cdot x_{t-l}\\&\text{Mr}:\quad \hat{y}_t^{\text{r}}\sim\sum_{l=1}^{L}a_l\cdot y_{t-l}\end{aligned}\right\}\tag{1}$$

采用F检验法判别Mf是否比Mr的预测更准确(更优)，为此构造F统计量

$$F=\frac{(\text{SSE}_{\text{r}}-\text{SSE}_{\text{f}})/L}{\text{SSE}_{\text{f}}/(T-2L-1)}\tag{2}$$

其中，预测误差平方和$\mathrm{SSE_r}=\sum_{t=1}^{T}(y_t-\hat{y}_t^{\mathrm{r}})^2$，$\mathrm{SSE_f}=\sum_{t=1}^{T}(y_t-\hat{y}_t^{\mathrm{f}})^2$。

在零假设(Mf的预测不显著优于Mr)下，若由时间序列数据计算得出的F值大于显著性水平α下的阈值$c^*_{L,T-2L-1}$，则拒绝上述零假设，判定X是引起Y的原因($X\rightarrow Y$)；否则，判定无因果关系$X\nrightarrow Y$。类似地可以判别Y是否是引起X的原因($Y\rightarrow X$)。

F界检测法[13]主要设计有关于时间的外循环迭代每个时刻t(t=1, 2,…, T)，内循环迭代时间区间[i,i+j](其中i=t, j=1, 2,…, T–t)：首先执行起步检测，即对起步区间[i,k](k–i+1的初值可设置为由时滞L限定的最小长度)内的时间序列采用上述Granger因果检测方法进行因果关系的检测；然后，计算扩展区Ut=[i, k+j](其中j=1, 2,…, T–k)内F统计量的上下界，使用该上下界判别Ut内的因果关系，若满足起步条件则k=k+j并执行起步检测；继续下一次迭代。最后对每个t，计算包含t的因果区间数量占全部区间数量的比值作为因果分，用因果分判别是否存在因果关系。F界检测法的主要贡献在于使用了已检测过区域的拟合模型和误差项，替代计算扩展区Ut内的拟合模型和误差项，来构建F统计量的上下界，节省了计算量。

3 时间序列因果发现的差异区域平衡方法

常规滑动窗方法的主要弱点是其性能对窗口宽度敏感，不同宽度的窗口有可能带来不同的性能，不合适的窗宽很可能导致低性能。性能改进的新思路是：将这种单一滑动窗口(单一区域)的因果检测方式改进为3个有差异区域(即当前滑动窗口、当前窗口扩展出的前向和后向区域)的因果检测方式；于是，当一个区域的数据体现的序列间关系有偏差，而其余2个区域的数据体现的序列间关系为正确时，以多数票赢方式综合3个区域的结果仍为正确。这样，与单一区域相比，3区域方式可降低性能对窗口宽度的敏感性。次要弱点是性能对滑动窗的移动步长也敏感。考虑到移动步长的主要作用是控制窗口移动的快慢和计算量，改进性能的思路是：将步长限制在较小的范围内，从而限制步长敏感性在一个较低的范围。

具体地设计与当前时间窗口Wt紧相邻的长wide/2的前向区域作为前向相邻窗口Ut；与窗口Wt紧相邻的长wide/4的后向区域作为后向相邻窗口Vt。在探索时间序列间关系的常规滑动窗方法的框架基础上，差异区域平衡方法的设计包括如下措施：

(1) 前向探索策略：计算窗口Wt内序列的波动程度Sw，并作为局部波动界；计算窗口Ut内序列的波动程度Su；若Su≤Sw，实施差异区域的平衡检测；若Su>Sw，前向区域出现序列间关系转换的可能性增大，则实施对称的差异平衡检测。

(2) 差异区域的平衡检测：对窗口Wt(长wide)内序列间关系检测一次，对窗口Wt与Vt的合并区域(长1.25 wide)进行一次关系检测，对窗口Wt与Ut的合并区域(长1.5 wide)内序列间关系检测一次。将这3次不同长度区域的检测结果进行综合作为当前窗口上的序列间关系，有利于降低单一长度区域时检测结果可能的偏差影响，提高因果检测的准确性和稳定性。

(3) 对称区域的平衡检测：采用上述平衡检测方案，但其中窗口Vt的长度改为wide/2，此时对窗口Wt而言，窗口Ut与Vt是前后对称的。当前窗口出现序列间关系的转换点时，这种对称方式体现了转换点前后检测区域的平衡性，有利于降低转换点区域的检测偏差。

(4) 多次检测结果的综合：多次检测可能出现不一致的结果，则由多数票赢的投票法来决策有无因果关系。

差异区域平衡法需要时间窗口内数据波动程度的衡量方法，为此设计如下方法。

定义1 (平均单间隔差分)设时间轴上在时刻t的滑动时间窗口Wt包含时间序列Y的m+1个元素$\{y_{t-m},\cdots,y_{t-1},y_t\}$，衡量窗口Wt中元素波动幅度的平均单间隔差分定义为

$$\mathrm{sf[Wt]}=\frac{1}{m}\sum_{k=0}^{m-1}|y_{t-k}-y_{t-k-1}| \tag{3}$$

定义2 (平均双间隔差分)设时间轴上在时刻t的滑动时间窗口Wt包含时间序列Y的$2m$+1个元素$\{y_{t-2m},\cdots,y_{t-1},y_t\}$，衡量窗口Wt中元素波动幅度的平均双间隔差分定义为

$$\mathrm{df[Wt]}=\frac{1}{2m}\sum_{k=0}^{m-1}[|y_{t-2k}-y_{t-2k-2}|+|y_{t-2k-1}-y_{t-2k-3}|] \tag{4}$$

差异区域平衡法的实现过程由3部分组成：主框架部分是滑动窗的迭代移动过程，即宽wide的滑动窗口Wt以步长step沿时间t方向移动；第2部分是在移动到当前窗口Wt时，依据前向探索策略检测时间序列X和Y之间的关系；第3部分是综合多次检测结果作为输出，即每个时间点所记录的多次检测结果(有或无关系)采用多数票赢的投票法。

设两个长度为T的时间序列X和Y，仅评估目标变量Y的波动界，降低波动界为0.95Sw以加强分

析。差异区域平衡法的实现步骤设计如下：

步骤 1　确定滑动时间窗口$W=[i,j]$的宽度wide、滑动步长step_slide，序列的自回归时滞参数L；固定参数有最大探索区域长度uwide=0.5wide，区域长度增量du=0.5uwide；初始参数$i=1$, $j=$wide, $j_tail=j$，时长增量dt=0，后向增量dback=0，检测结果$Q=\{q_k=0\}, k=1:T$；

步骤 2　窗口$W=[i, j]$上的波动界计算：按前文定义计算Y的平均单、双间隔差分sf[i, j]与df[i, j]作为波动界，dback=du；

步骤 3　前向探索波动程度：dt=dt+du, $j_tail=j$+dt，前向区域U=[j, j_tail]，计算U上Y的sf[j, j_tail]与df[j, j_tail]；若sf[j, j_tail]$>0.95\times$sf[i, j]或df[j, j_tail]$>0.95\times$df[i, j]，则令dback=uwide，转下一步骤；若dt≤du，重复本步骤；

步骤 4　平衡检测之后向检测：$i_head=\max(0, i-\text{dback})$，对后向合并区域[$i_head$, j]内时间序列X和Y采用Granger因果检测方法进行检测；若有因果关系$X\rightarrow Y$或$X\leftarrow Y$，记录$q_k=q_k+1, k=i_head:j$；若无因果关系，记录$q_k=q_k-1, k=i_head:j$; dt=0；

步骤 5　平衡检测之前向检测(包括窗口W及前向合并区域$W+U$)：$j_tail=j$+dt，对区域[i, j_tail]内时间序列X和Y采用Granger因果检测方法进行检测；若有因果关系$X\rightarrow Y$或$X\leftarrow Y$，记录$q_k=q_k+1, k=i:j_tail$；若无因果关系，$q_k=q_k-1, k=i:j_tail$；若dt<uwide，则令dt=uwide，重复本步骤；

步骤 6　窗口W移动：若i+step_slide+wide+uwide≤T，令$i=i$+step_slide, $j=i$+wide, dt=0，转步骤2；否则，若$j<T$，令$j=T$, $i=\min(j-\text{wide}, i+\text{step_slide})$, dt=0, uwide=0，转步骤5；

步骤 7　输出结果：Q中$q_k>0$的区域为有因果关系，$q_k\le 0$的区域为无因果关系；停止。

现在分析本文方法比现有的常规滑动窗方法、F界检测法和转折点方法具有更好的性能稳定性。考虑图1所示的一般情况，设很少数据点的时间区域上的序列间关系是无意义的；设两条时间序列X和Y在时间区域[t_1, t_2]具有近似线性关系，而在区域[t_2, t_7]没有线性关系，但在区域[t_4, t_5]出现干扰性波动(有线性关系的假象)。

对于常规滑动窗方法，由于区域[t_4, t_5]中体现出近似线性关系的数据占比大约为60%，该方法的窗宽小于该区域宽度时的Granger因果检测结果是存在因果关系的概率大，即出现误判的概率大。而本文方法虽然在区域[t_4, t_5]出现误判的概率大，但在区域[t_4, t_6]和[t_3, t_5]体现出近似线性关系的数据占比明显小于50%，使得Granger因果检测结果是存在因果关系的概率小，即出现误判的概率小；于是，3个差异区域[t_4, t_5], [t_4, t_6]和[t_3, t_5]的多数票赢带来正确结果的概率大，即本文方法比常规滑动窗方法具有更稳定的性能。

F界检测法的遍历方式偏重全局信息，局部信息倾向于被抑制或弱化。该方法出现弱点的情形是，图1中当区域[t_1, t_2]相对于[t_2, t_7]比较短时(情形1)，短区域[t_1, t_2]在全区域上占比小，则其(以区间数量占比为基础的)因果分会比较小，因大多数遍历区域属于区域[t_2, t_7]；相反地，当区域[t_1, t_2]相对于[t_2, t_7]比较长时(情形2)，区域[t_2, t_7]上的因果分比较大，因大多数遍历区域属于区域[t_1, t_2]。这两种情形会带来因果分判别准则设计的矛盾性，即取大的因果分阈值会使情形1的区域[t_1, t_2]出现误判的概率大，取小的因果分阈值会使情形2的区域[t_2, t_7]出现误判的概率大；于是，上述复杂情形容易导致F界检测法性能下降。而本文方法以窗口局部信息为主，不受全局信息干扰，故比F界检测法具有更稳定的性能。

转折点方法出现弱点的情形是，当序列数据波动多时找出的转折点也多，但时间序列波动的转折点也可能不是序列间关系的转折点，例如图1中时刻t_5位置就不是序列间关系的转折点，这种误识别会使该方法性能下降。而本文方法不涉及序列波动的转折点，上述情形不影响本文方法的性能，故本文方法比转折点方法具有更稳定的性能。

4　实验结果

本节对差异区域平衡法(简记为平衡)与常规滑动窗方法(简记为常规)、F界检测法(简记为F界)、转折点检测法(简记为转折)进行实验对比，各方法的程序均采用如下相同的参数和设置：运行环境为pycharm和python 3.8.3；主框架为滑动窗移动过程，相邻区域如Ut与Vt的长度不超过窗口Wt长度的1/2；时滞参数L给出约束条件，即滑动窗口或检测区域的长度均不小于$2L+2$；F检验的显著性水平为0.05。对于转折法，奇异谱分析的嵌入维度为6，对前3条奇异谱进行距离计算，距离计算所用

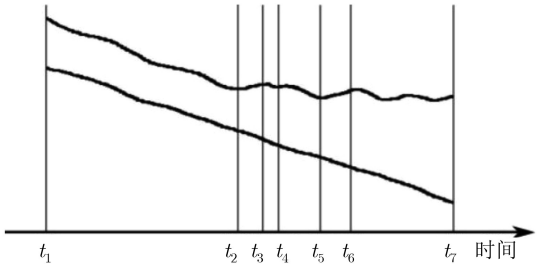


图 1 两条序列的有近似线性关系区域[t_1,t_2]和无线性关系区域[t_2, t_7]

的序列长度为窗口Wt长度的1/2，转折点的判别阈值为2倍的历史平均距离(任一条奇异谱序列满足即可)，相邻的2个转折点确定一个检测区域。对于F界检测法，以步长1移动的同时遍历包含时刻t的所有不同长度的时间区域；判别阈值取最高因果分的90%，因果分大于该阈值的区域为有因果关系。

性能评价指标采用正确率，即统计各个时间点上算法的输出结果与已知因果关系相同的数量(m)，计算该数量占总时间点数量(T)的百分比($100\%\times m/T$)。

模拟数据集参照文献[13]产生长1000的平稳时间序列X和Y，时滞参数L=2，有关系区间[s^*, t^*]=[450, 650], X的产生公式为$x_t=1+\sin(0.08t)+\varepsilon_t$，其中$\varepsilon_t$为零均值与方差$\sigma^2$的高斯噪声。每次模拟实验都产生一个随机模拟数据集，包括噪声ε_t随机产生，序列Y的初值y_0和y_1从0～0.1的均匀分布中随机采样，关系区间的边界s^*和t^*分别添加区间[–20, 20]中的随机整数，使得s^*和t^*每次实验不相同。使用不同的窗口宽度wide和步长step_slide组合(但不适用于F界检测法)，对各种方法重复进行50次性能实验，获得的平均结果列于表1中。

真实数据集Dropoff-tweet来自文献[13]的图4，是2012年10月在纽约会议中心下客的出租车数量与观众评论帖数量随时间变化的时间序列，样本数量为745(t=1～745)。在纽约会议中心有喜剧演出的11～14日，下客的出租车数量(Dropoff序列X)是引起该演出的评论帖数量(Tweet序列Y)变化的原因。对序列X和Y的样本计算平方根作为数据预处理；时滞参数L=5。这样，时滞对齐后的序列X和Y在演出区间[248, 334]具有近似线性关系(因果关系)，其中每天的夜眠时间(凌晨1～7点)不存在因果关系。实验任务是检测31天中在何时段时间序列之间存在因果关系。考虑一天24小时内序列X和Y之间的影响关系，采用不同的窗口宽度wide和步长step_slide的组合(但不适用于F界检测法)，其中F界检测法遍历的最大区间长度为72(与文献[13]一致)，实验结果列于表2中。

真实数据集Tweet-pickup来自文献[13]的图4，是2012年10月在纽约会议中心上客的出租车数量与观众评论帖数量随时间变化的时间序列。演出结束前后时间，关于演出的评论帖数量(Tweet序列X)是引起上客的出租车数量(Pickup序列Y)变化的原因。其余信息与数据集Dropoff-tweet相同，实验结果列于表3中。

真实数据集Fish-school[15]是未训练的鱼(游动方向序列Y)跟随被训练的鱼(游动方向序列X)游向喂食点的时间序列集中第2个子集，样本数量为596(t=1～596)。实验任务是检测该时间序列之间在何时段存在因果关系。数据做如下预处理：对序列X计算每20个(类似地对Y是每10个)连续的序列样本的均值(滚动均值)，获得在t=20～596时的平滑样本，而t=1～20时的样本取t=20时的值；然后对平滑样本计算平方根。时滞参数L=65。这样，在鱼游向喂食点期间t=300～479，时滞对齐后的序列X和Y之间具有近似线性关系(因果关系)。考虑最小窗宽需大于132，实验采用不同的窗口宽度wide和步长step_slide组合(但不适用于F界检测法)，实验结果列于表4中。

真实数据集Baboon-troop[15]是一队狒狒(移动方向序列Y)跟随狒狒ID3(移动方向序列X)移动的时间序列集中第1个子集，样本数量为599(t=1～

表 1 不同方法在模拟数据集上发掘因果关系的正确率(%)

窗口宽度	滑动步长	噪声方差0.01				噪声方差0.2				噪声方差0.5			
		常规	F界	转折	平衡	常规	F界	转折	平衡	常规	F界	转折	平衡
20	5	91.32	88.18	90.99	**95.40**	80.82	**94.56**	58	84.13	80.57	**91.00**	50.29	82.52
	10	88.98	88.18	91.91	**95.06**	80.25	**94.56**	53.39	83.33	80.15	**91.00**	46.33	82.54
	15	86.78	88.18	90.00	**94.43**	78.1	**94.56**	44.86	83.26	78.52	**91.00**	46.22	82.03
	20	83.65	88.18	89.17	**92.77**	77.05	**94.56**	45.4	82.15	77.98	**91.00**	45.48	81.87
30	5	95.85	88.18	90.61	**95.87**	85.69	**94.56**	60.4	92.31	83.63	**91.00**	47.37	87.78
	10	94.43	88.18	90.15	**95.49**	84.95	**94.56**	53.01	91.95	83.69	**91.00**	46.4	87.41
	15	94.07	88.18	90.85	**94.91**	83.22	**94.56**	51.62	91.65	81.93	**91.00**	45.98	86.84
	20	92.96	88.18	90.95	**95.57**	81.62	**94.56**	53.21	91.37	81.18	**91.00**	46.38	86.96
40	5	**95.56**	88.18	91.83	94.85	92.46	94.56	63.52	**94.65**	87.62	91.00	46.43	**92.17**
	10	**95.31**	88.18	90.07	94.87	91.05	**94.56**	58.89	94.27	87.08	91.00	45.38	**92.22**
	15	**94.95**	88.18	90.65	94.77	90.62	**94.56**	58.77	94.55	86.87	91.00	46.39	**91.76**
	20	**94.59**	88.18	89.9	94.31	89.5	**94.56**	52.31	93.93	85.86	91.00	45.80	**91.03**

表 2　在数据集Dropoff-tweet上发掘因果关系的正确率(%)

窗口宽度	滑动步长	常规滑动窗	F界检测法	转折点法	差异平衡法
12	4	91.95	92.62	**93.56**	93.42
	8	90.87	92.62	93.56	**93.83**
	12	89.26	92.62	**94.36**	90.20
18	4	94.09	92.62	94.90	**95.30**
	8	94.36	92.62	94.36	**94.77**
	12	92.21	92.62	94.36	**94.77**
24	4	94.36	92.62	94.09	**96.51**
	8	97.05	92.62	96.24	**96.78**
	12	91.95	92.62	91.41	**95.70**

表 3　在数据集Tweet-pickup上发掘因果关系的正确率(%)

窗口宽度	滑动步长	常规滑动窗	F界检测法	转折点法	差异平衡法
12	4	90.87	**94.90**	90.87	93.29
	8	91.95	**94.90**	94.09	94.09
	12	92.48	**94.90**	94.09	91.01
18	4	92.48	**94.90**	88.19	**94.90**
	8	92.48	**94.90**	93.83	94.90
	12	93.02	**94.90**	94.09	93.02
24	4	92.75	94.90	91.41	**95.44**
	8	93.29	94.90	82.82	**95.97**
	12	92.21	94.90	95.44	**95.44**

表 4　在数据集Fish-school上发掘因果关系的正确率(%)

窗口宽度	滑动步长	常规滑动窗	F界检测法	转折点法	差异平衡法
140	10	89.60	54.19	69.80	**90.1**
	20	86.24	54.19	69.80	**93.29**
	30	91.28	54.19	69.80	**95.64**
150	10	**89.60**	54.19	69.80	84.90
	20	83.22	54.19	69.80	**91.28**
	30	89.93	54.19	69.80	**99.66**
160	10	83.22	54.19	69.80	**86.91**
	20	69.80	54.19	69.80	**93.62**
	30	81.54	54.19	69.80	**92.95**

表 5　在数据集Baboon-troop上发掘因果关系的正确率(%)

窗口宽度	滑动步长	常规滑动窗	F界检测法	转折点法	差异平衡法
110	10	**80.63**	35.39	59.10	**80.63**
	20	70.62	35.39	59.10	**82.30**
	30	70.62	35.39	59.10	**80.47**
120	10	80.63	35.39	59.10	**81.64**
	20	62.27	35.39	59.10	**83.31**
	30	63.94	35.39	59.10	**83.97**
130	10	80.80	35.39	59.10	**82.30**
	20	75.79	35.39	59.10	**83.97**
	30	82.30	35.39	59.10	**82.97**

599)。实验任务是检测该时间序列之间在何时段存在因果关系。对序列X和Y的数据均做如下预处理：计算每25个连续的序列样本的均值(滚动均值)，获得在$t=25\sim596$时的平滑样本，而$t=1\sim24$时的样本取$t=25$时的值。时滞参数L=12。这样，在一队狒狒跟随狒狒ID3移动期间t=92～445，时滞对齐后的序列X和Y之间具有近似线性关系(因果关系)。考虑是近似线性关系，在较短时间区间上序列间关系不稳定，故取较大窗宽进行实验；实验采用不同的窗口宽度wide和步长step_slide组合(但不适用于F界检测法)，实验结果列于表5中。

从上面各表中可以看出，差异区域平衡法成功地揭示出时间序列之间随时间变化的因果关系。在数据集Dropoff-tweet,Tweet-pickup和模拟数据集上，平衡方法的性能与对比方法相当，其原因是：出现因果关系的区域刚好位于全部时间区域的中部，且序列X和Y出现显著增大/变化的值(即体现因果关系的信号强)，而在其他时间区域序列X的值为0或接近0，这种区分性极其显著的数据集使得各方法容易获得好的性能。在数据集Fish-school和Baboon-troop上，绝大部分情况下平衡方法的正确率都较大幅度高于对比方法；对比方法性能较低的原因是，数据集出现上节提及的因果区间长于非因果区间、序列数据波动较多这样导致对比方法性能下降的情形。虽然在少量情况下平衡方法的正确率略低于对比方法，但在高正确率且性能稳定的综合性能上，差异区域平衡法优于对比方法。

实验中还获得关于参数的观察体会：对于转折点方法，从表1可以看出，其性能随着噪声的增大下降很多，即容易受数据波动的影响；该方法更适合时间序列波动的波峰波谷差别显著且与序列间关系转换对应的情况；当序列波动较小时应当设置较小的判别阈值，但小阈值很可能找出过多的波峰波谷(转折点)，而设置大阈值则可能丢失转折点；因此，阈值参数本身就是个矛盾体。对于F界检测法，实验中使用的阈值是依据多次实验经验选取的最好阈值，如何更好地设置因果分阈值仍需进一步的研究。对于常规滑动窗方法，文献[12]建议窗宽不超过20，而上文的实验结果表1—表3显示，窗宽超过20反而出现更好的性能。因此，如何调整参数以提升性能仍需要更深入的研究。

上述实验结果显现出窗口宽度和滑动步长对正确率有一定的影响。故而，需要根据实际问题相对合理地选取窗口宽度和滑动步长。在相对合理的窗宽和步长下，正确率高和性能稳定都是我们选用方法的考量指标。因此，从正确率和稳定性这两方面的综合角度考量，差异区域平衡法胜过对比方法。

5 结束语

在常规滑动窗方法和Granger因果检测的基础上，本文方法新设计了3个差异区域替代单一滑动窗口，新设计了3个差异区域的平衡检测方式，设计了多次检测结果不一致时按多数票赢的投票法确定最终是否存在因果关系的方法。这种差异区域平衡策略有利于降低单一区域检测结果可能的偏差影响，有利于降低方法性能对窗宽的敏感性，提高因果关系检测的准确性和稳定性。滑动窗方法的窗口宽度和滑动步长之间是否存在某种影响性能的关联性值得进一步研究。

参 考 文 献

[1] XIE Feng, CAI Ruichu, ZENG Yan, *et al*. An efficient entropy-based causal discovery method for linear structural equation models with IID noise variables[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 31(5): 1667–1680. doi: 10.1109/TNNLS.2019.2921613.

[2] YANG Jing, GUO Xiaoxue, AN Ning, *et al*. Streaming feature-based causal structure learning algorithm with symmetrical uncertainty[J]. *Information Sciences*, 2018, 467: 708–724. doi: 10.1016/j.ins.2018.04.076.

[3] 任伟杰, 韩敏. 多元时间序列因果关系分析研究综述[J/OL]. 自动化学报, https://doi.org/10.16383/j.aas.c180189, 2019.
REN Weijie and HAN Min. Survey on causality analysis of multivariate time series[J/OL]. *Acta Automatica Sinica*, https://doi.org/10.16383/j.aas.c180189, 2019.

[4] HUANG Biwei, ZHANG Kun, GONG Mingming, *et al*. Causal discovery and forecasting in nonstationary environments with state-space models[C]. The 36th International Conference on Machine Learning, Long Beach, USA, 2019.

[5] DU Sizhen, SONG Guojie, HAN Lei, *et al*. Temporal causal inference with time lag[J]. *Neural Computation*, 2018, 30(1): 271–291. doi: 10.1162/neco_a_01028.

[6] GRANGER C W J. Investigating causal relations by econometric models and cross-spectral methods[J]. *Econometrica*, 1969, 37(3): 424–438. doi: 10.2307/1912791.

[7] ORJUELA-CAÑÓN A D, CERQUERA A, FREUND J A, *et al*. Sleep apnea: Tracking effects of a first session of CPAP therapy by means of Granger causality[J]. *Computer Methods and Programs in Biomedicine*, 2020, 187: 105235. doi: 10.1016/j.cmpb.2019.105235.

[8] 范立夫, 赵善学, 张永军. 信贷结构和产业结构的相互影响研究——基于异质面板数据的格兰杰因果检验[J]. 宏观经济研究, 2019(6): 73–82. doi: 10.16304/j.cnki.11-3952/f.2019.06.007.
FAN Lifu, ZHAO Shanxue, and ZHANG Yongjun. Research on the interaction between credit structure and industrial structure——Granger causality test based on heterogeneous panel data[J]. *Scientific Management Research*, 2019(6): 73–82. doi: 10.16304/j.cnki.11-3952/f.2019.06.007.

[9] 李永立, 吴冲. 基于多变量的Granger因果检验方法[J]. 数理统计与管理, 2014, 33(1): 50–58. doi: 10.13860/j.cnki.sltj.2014.01.003.
LI Yongli and WU Chong. The Granger causality test method based on the multiple variables[J]. *Journal of Applied Statistics and Management*, 2014, 33(1): 50–58. doi: 10.13860/j.cnki.sltj.2014.01.003.

[10] REN Weijie, LI Baisong, and HAN Min. A novel Granger causality method based on HSIC-Lasso for revealing nonlinear relationship between multivariate time series[J]. *Physica A*: *Statistical Mechanics and its Applications*, 2020, 541: 123245. doi: 10.1016/j.physa.2019.123245.

[11] FINKLE J D, WU J J, and BAGHERI N. Windowed Granger causal inference strategy improves discovery of gene regulatory networks[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2018, 115(9): 2252–2257. doi: 10.1073/pnas.1710936115.

[12] CHANG T, TSAI S L, and HAGA K Y A. Uncovering the interrelationship between the U. S. stock and housing markets: A bootstrap rolling window Granger causality approach[J]. *Applied Economics*, 2017, 49(58): 5841–5848. doi: 10.1080/00036846.2017.1346365.

[13] LI Zhenhui, ZHENG Guanjie, AGARWAL A, *et al*. Discovery of causal time intervals[C]. 2017 SIAM International Conference on Data Mining, Houston, USA, 2017: 804–812.

[14] MASNADI-SHIRAZI M, MAURYA M R, PAO G, *et al*. Time varying causal network reconstruction of a mouse cell cycle[J]. *BMC Bioinformatics*, 2019, 20: 294. doi: 10.1186/s12859-019-2895-1.

[15] AMORNBUNCHORNVEJ C, ZHELEVA E, and BERGER-WOLF T Y. Variable-lag granger causality for time series analysis[C]. 2019 IEEE International Conference on Data Science and Advanced Analytics, Washington, USA, 2019.

王开军：男，1965年生，副教授，研究方向为机器学习和数据挖掘.
曾元鹏：男，1995年生，硕士生，研究方向为模式识别和数据挖掘.

责任编辑：余 蓉