Info Vis 2022 論文書面報告：

Compass: Towards Better Causal Analysis of Urban Time Series

Zikun Deng, Di Weng, Xiao Xie, Jie Bao, Yu Zheng, Mingliang Xu, Wei Chen, and Yingcai Wu
IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS, VOL. 28, NO. 1, JANUARY 2022

第十二組 曹昱維(110753201)

ABSTRACT

- This paper presents Compass, a novel visual analytics approach for in-depth analyses of the dynamic causality in urban time series.
- This approach aims to enable analysts to obtain and understand correct urban causalities in dynamic environments.

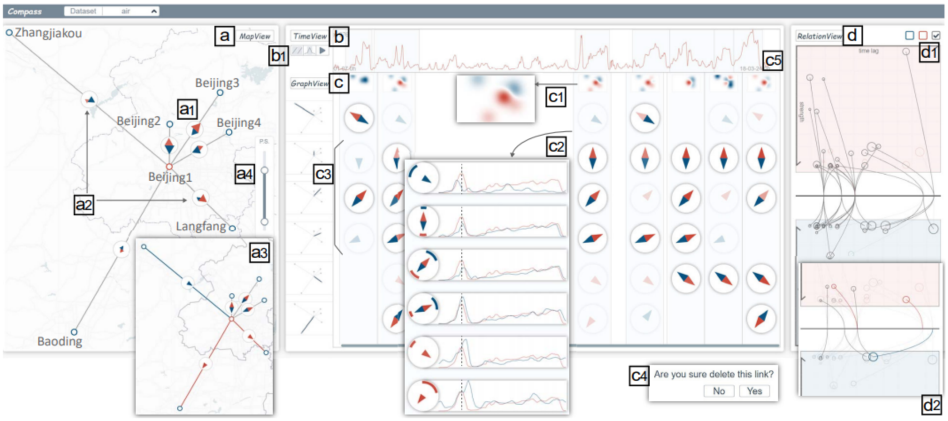


Fig. 1. The interface of Compass. (a) The map view enables users to select a sensor to start analysis and reason causal graphs within a spatial context. (b) The time view presents the time series of the selected sensor and time partitioning results. (c) The graph view visualizes the dynamic causal graphs detected by a causal detection framework along the same timeline of the time view. (d) The relation view presents the causal relations involved in the causal graphs with a multi-dimensional visualization.

INTRODUCTION

- To develop Compass approach, we identify and address three challenges:
  - Detecting fine-grained urban causality

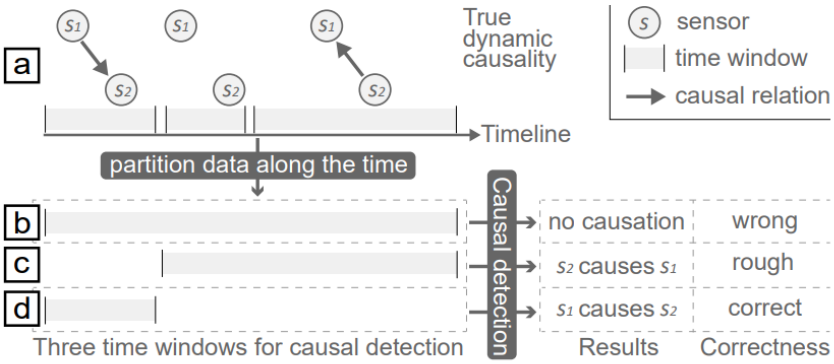


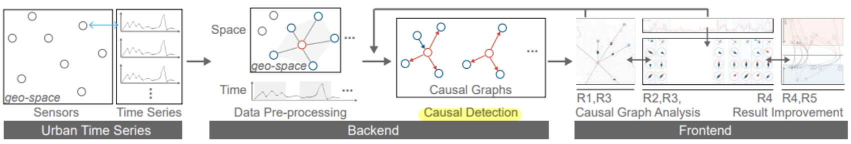
Fig. 2. Motivation illustration. (a) Real-world causal relations between sensors can change over time, shown by the different relations in three time windows. (b, c) Causal detection in large time windows may produce wrong or rough results. (d) Although the result is correct, further interpretation and verification are still required for informed policy making.

  - Interpreting dynamic causal relations
  - Unveiling suspicious causal relations

SYSTEM OVERVIEW

Workflow

- WP: Pre-processing data.
  - A target sensor is specified as **ego** first, and then its neighbors that have potential causality will be retrieved
  - the time is properly **partitioned** into windows to reveal dynamic causalities.
- WA: Analyzing causal graphs.
  - multiple causal graphs centered on the ego sensor can be obtained in the partitioned time windows
- WI: Improving causal detection results.
  - design interactive visualizations tailored for our problem to assist users in discovering and **modifying the incorrect causalities**.



Requirement Analysis

- R1 Summarize causal graphs across the time (WA).
- R2 Explore causal graphs along the time (WA).
- R3 Learn influence propagation via causal graphs (WA).
- R4 Interpret and validate causal relations (WA, WI).
- R5 Modify incorrect causal relations (WI).

MODEL

Granger causality

- vector autoregressive (VAR) model
  - In VAR, the current state of a system can be **predicted by the past K states** in different time series across the system

$$\begin{cases} \mathbf{v}_t^1 = \mu_1 + \sum_{n=1}^{N} \sum_{k=1}^{K} \omega_{1,n,k}\, \mathbf{v}_{t-k}^n + u_{1,t} \\ \mathbf{v}_t^2 = \mu_2 + \sum_{n=1}^{N} \sum_{k=1}^{K} \omega_{2,n,k}\, \mathbf{v}_{t-k}^n + u_{2,t} \\ \dots \\ \mathbf{v}_t^N = \mu_N + \sum_{n=1}^{N} \sum_{k=1}^{K} \omega_{N,n,k}\, \mathbf{v}_{t-k}^n + u_{N,t} \end{cases}$$

  - \(v^n_t\) : time t ; n-th variable(*sensor*)
- Causality Test.
  - Testing “x → y” (x causes y) is based on the following two regression equations:

$$\mathbf{v}_t^y = \mu_y + \sum_{n=1}^{N} \sum_{k=1}^{K} \omega_{y,n,k}\, \mathbf{v}_{t-k}^n + u_{y,t} \qquad (1)$$

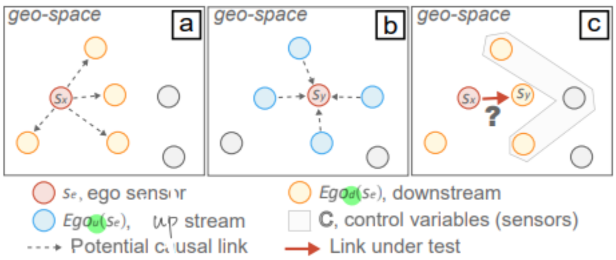
$$\mathbf{v}_t^y = \mu_y + \sum_{n=1, n \neq x}^{N} \sum_{k=1}^{K} \omega_{y,n,k}\, \mathbf{v}_{t-k}^n + u_{y,t} \qquad (2)$$

    - All variables except for x and y are viewed as control variables **C**.
  - The causal strength
    - Use F Test to estimate, measured by (0.05−p-value)/0.05.

$$F = \frac{(SSR_{(1)} - SSR_{(2)})/K}{SSR_{(2)}/(M - KN)}.$$

Data Pre-processing

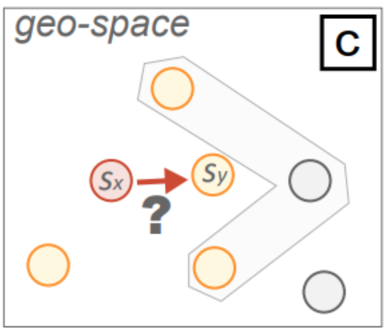
- Indexing neighbors.
  - For every sensor, we index its neighbors with potential causality and categorize them into downstream and upstream neighbors based on actual applications.



- Partitioning time.
  - For periodic time series,they can be directly partitioned by their period
  - The peaks of the time series will be extracted automatically, and thereby the time windows can be identified based on these peaks

Causal Link Testing

- Testing the causal links between in ego sensor and each of its neighbor sensors in every partitioned time window.
- When testing the link $s_x \rightarrow s_y$, $Ego_u(s_y) \backslash \{s_x\}$ are naturally considered control variables **C**

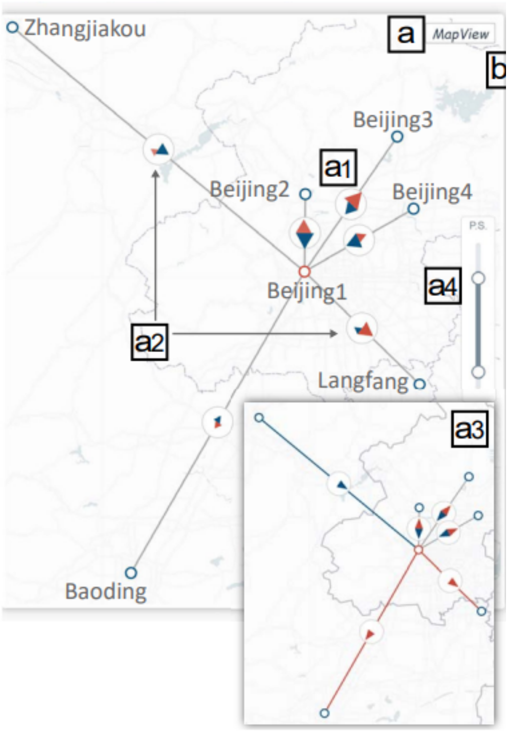


- K can be determined based on actual applications and the spatial distance between sensors.

VISUAL DESIGN

Map View

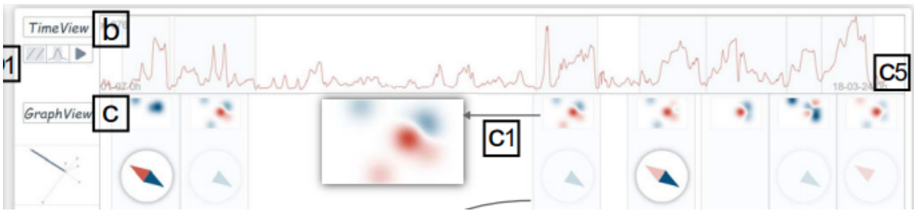
- The map view provides a spatial context for selecting the ego sensors of interests and reasoning causal graphs (R1 and R3).



- The causal directions across multiple graphs are first aggregated according to the edge. The aggregation is then encoded with a revisited compass glyph
- The arrow's size encodes the frequency of the causal links with this direction. An offset between the two opposite arrows can be observed if bi-directional relations exist. The overlapping part of the two arrows encodes the frequency of bi-directional relations.

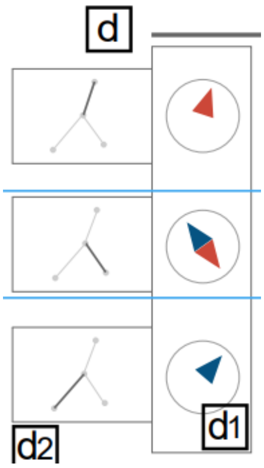
Time View

- The time view shows the time series of selected sensors and the partitioned time windows. It also serves as a timeline for the graph view.

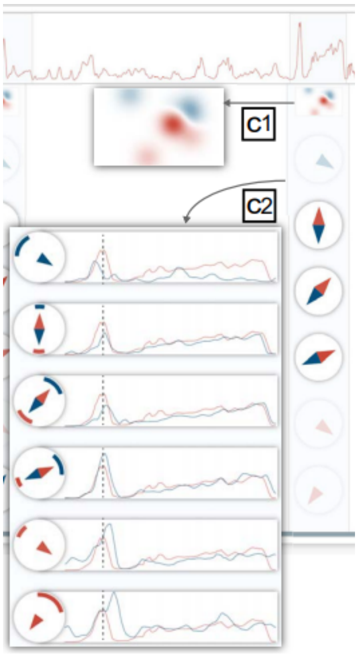


Graph View

- The graph view displays the causal graphs detected in each window along the timeline.
- In this view, users can explore the detected dynamic causal graphs along the time (R2) and interpret each graph with its causal relations (R3 and R4) from a spatiotemporal perspective.
- Visualizing single causal graph (R3)



- Visualizing the causal relations of a graph (R4).



- Visualizing multiple causal graphs(R2).
    - The causal directions across multiple graphs are first aggregated according to the edge. The aggregation is then encoded with a revisited compass glyph
    - The arrow's size encodes the frequency of the causal links with this direction. An offset between the two opposite arrows can be observed if bi-directional relations exist. The overlapping part of the two arrows encodes the frequency of bi-directional relations.

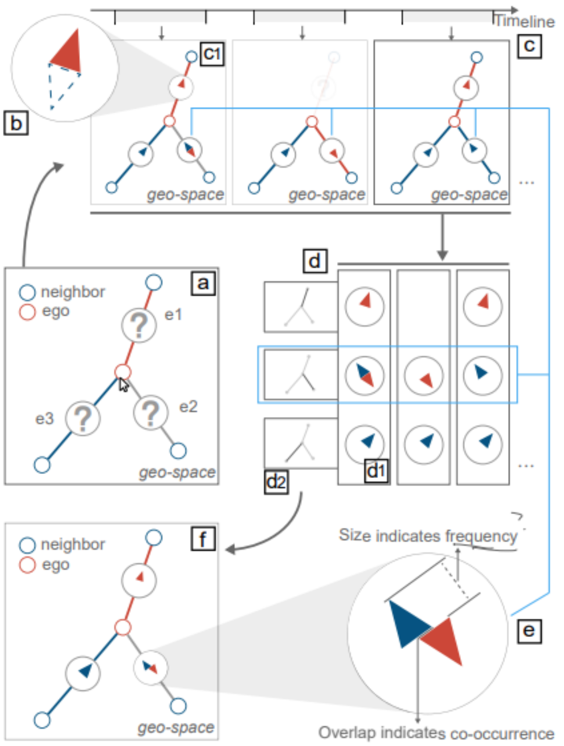


Fig. 5. Visual encodings for causal graphs. (a) An illustrative spatial ego-graph comprises three edges (i.e., e1, e2, and e3) with three different color styles. (b) A compass glyph shows the causal relation where only the causal link from the ego sensor exists. (c) Multiple causal graphs are detected in different time windows and depicted as spatial causal graphs on maps. (d) Graph bands are aligned along the timeline as a compact visualization of dynamic causal graph. (e) A revisited compass glyph summaries the compass glyphs of the edge e2. (f) A spatiotemporal causal graph summaries the spatial causal graphs in (c).

Relation View

- The relation view presents every causal relation from multiple dimensions.
- Users can interpret causal relations further (R4), discover spurious ones, and improve causal detection results (R5).
- Showing multidimensional details(R4).

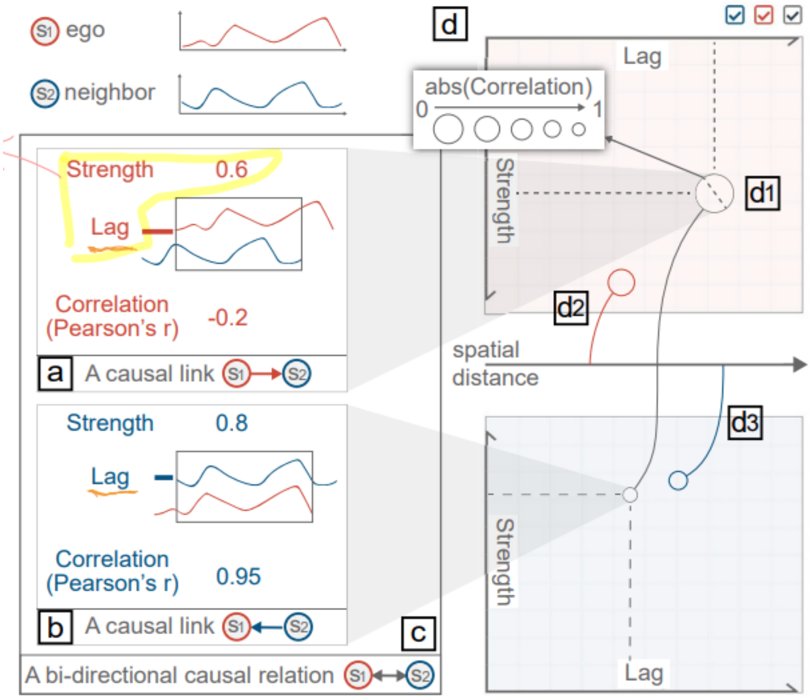


Fig. 6. A causal relation (left) and the relation view for visualizing causal relations (right). (a) A causal link may be incorrect because the Pearson's r is near 0. (b) A causal link opposite to (a) has a large Pearson's r. (c) A bi-directional causal relation comprises two causal links with opposite directions. (d) The relation view adopts a multidimensional visualization and presents (d1, d2, and d3) three illustrative causal relations.

  - These circles also connect to the middle axis with curves according to the two involved sensors' spatial distance.
  - If two links are comprised by the same causal relation, their curves are regarded as one and correspond to the relation (Fig. 6d-1). We call these curves relation curves.
- Unveiling suspicious causalities(R5).
  - Bi-directional relations with a large spatial distance.
  - Causal links with low Pearson's r, i.e., large circles.

Expert Interviews

- Conducted informal interviews with three experts (EA, EB, and EC) and collected their feedback
  - EA and EB are urban computing experts
  - EC is a researcher with expertise in causal analyses.
  - The experts aimed to disclose how did air pollution influence Beijing and its surrounding areas.
- Visual design.
  - All three experts agreed that the visual designs of Compass were easy to learn and understand.
  - They praised the compass glyph, "*it is interesting and intuitive to indicate causal directions because the purpose of a compass is to show directions.*"
  - EA and EB can quickly understand the graph view although they have not seen dynamic graph visualizations before because "*the causal relations in the graph view can be clearly associated with the time and space.*"
- Usability and improvements.
  - All experts confirmed the system usability.
  - EA commented,"*Compass allows us to obtain fine-grained and dynamic urban causalities, which cannot be supported before.*"

CONCLUSION

- This study presents a novel visual analytics approach that assists analysts in detecting and analyzing dynamic causalities in urban domains.
- Compass facilitates analysts to interpret dynamic causal relations and improve causal detection results.