

DÉCOUVERTE DE CAUSALITÉ POUR SÉRIES TEMPORELLES EN PRÉSENCE DE CAUSES CACHÉES

Antonin Arzac ^{†,1} & Aurore Lomet ^{‡,2} & Jean-Philippe Poli ^{†,3}

[†] *Université Paris-Saclay, CEA LIST, F-91120, Palaiseau, France*

[‡] *Université Paris-Saclay, CEA, Service de Génie Logiciel pour la Simulation, 91191, Gif-sur-Yvette, France.*

¹ *antonin.arsac@cea.fr*, ² *aurore.lomet@cea.fr*, ³ *jean-philippe.poli@cea.fr*

Résumé. La recherche de causalité vise à définir les relations de causes à effets entre différentes variables à partir de données. Cet apprentissage peut s'avérer difficile notamment lorsque les observations d'un système complexe sont incomplètes et que l'intégralité des causes d'un phénomène ne sont pas observées. A cette complexité, le cadre spécifique des séries temporelles ajoute la nécessité de considérer des structures de dépendance telles que l'auto-corrélation. Le travail présenté dans ce papier se concentre sur l'apprentissage d'un graphe causal à partir de données de séries temporelles issues d'un système complexe en présence de causes latentes. Dans ce cadre, nous proposons une méthode permettant d'identifier des liens entre les séries, qu'ils soient linéaires ou non-linéaires et possiblement décalés dans le temps, pour construire un graphe causal. Pour ce faire, nous combinons un algorithme de découverte causale et une mesure basée sur la théorie de l'information. L'approche est évaluée sur un jeu de données simulées issues de la littérature, démontrant sa pertinence dans une étude comparative.

Mots-clés. Découverte causale, séries temporelles, théorie de l'information, variables latentes

Abstract. Causality aims to define cause-and-effect relationships between different variables based on data. Learning causality can be challenging, particularly when observations of a complex system are incomplete and all the causes of a phenomenon are not observed. To this complexity, the specific framework of time series data adds the necessity to consider dependency structures such as auto-correlation. The work presented in this paper focuses on learning a causal graph from time series data generated by a complex system in the presence of latent causes. In this context, we propose a method to identify relationships between the series, whether linear or non-linear and possibly time-lagged, to construct a causal graph. To achieve this, we combine a causal discovery algorithm with an information-theoretic measure. The approach is evaluated on simulated data from the literature, demonstrating its relevance in a comparative study.

Keywords. Causal discovery, time series, Information theory, latent variables

1 Introduction

L'étude de systèmes complexes évoluant dans le temps peut requérir l'analyse de séries temporelles. Ces données sont présentes dans de nombreux domaines comme la médecine

(au travers d’EEG par exemple), les analyses de marché [Brodersen et al., 2015] ou encore dans l’étude du climat [Runge et al., 2019a]. Explorer les relations de causes à effets entre différentes variables temporelles permet de comprendre au mieux ces phénomènes. En effet, simplement observer des corrélations décalées dans le temps peut s’avérer insuffisant : ces dernières peuvent provenir de variables cachées (variables latentes) et peuvent être très faibles dans le cas de systèmes non-linéaires bruités. De récents développements dans le domaine de la causalité, comme ceux de [Pearl, 2009], permettent la distinction entre fausse corrélation et relation causale, au travers des graphes dirigés acycliques (DAG). Ainsi, l’objectif de ce travail est de construire un graphe causal à partir de données de séries temporelles multi-variées en présence de variables non-observées.

Dans les graphes causaux, chaque noeud correspond à une variable et la présence d’une arête orientée entre deux noeuds représente la relation d’une cause vers son effet. Une absence d’arête entre deux noeuds se traduit par une indépendance ou une indépendance conditionnelle dans les données [Spirtes et al., 2000]. Cependant, un grand nombre de modèles et de méthodes reposent sur des hypothèses souvent peu réalistes comme la linéarité, une distribution *a priori* des données ou encore de suffisance causale, indiquant que toutes les causes de chaque effet sont observées.

Pour répondre à ces problèmes, les travaux présentés dans ce papier visent à inférer les relations causales entre des séries temporelles avec peu d’hypothèses. Pour ce faire, nous combinons un algorithme de découverte de causalité, l’algorithme FCI [Spirtes et al., 2000, Zhang, 2008] avec une mesure d’indépendance conditionnelle, la *Partial Mutual Information from Mixed Embedding* (PMIME) [Kugiumtzis, 2013]. L’algorithme FCI, couplé à un test d’indépendance conditionnelle, permet de construire un graphe causal en prenant en compte l’existence de variables non-observées. La mesure non-paramétrique PMIME, basée sur la théorie de l’information, développée pour les séries temporelles, considère l’auto-corrélation au sein des variables.

Dans cet article, nous introduisons en partie II les notions nécessaires pour traiter la causalité au sein des séries temporelles. La partie III présente la méthode que nous proposons. Enfin, nous réalisons des expérimentations dans la partie IV avant d’en discuter les résultats et de conclure en partie V.

2 Découverte causale et séries temporelles

2.1 Modèles de graphe

La causalité entre séries temporelles peut être représentée par des réseaux causaux bayésiens, une classe de modèles de graphe permettant une représentation probabiliste de variables aléatoires.

Le principe de priorité temporelle, qui stipule qu’une cause précède ses effets, induit l’asymétrie de la causalité dans le temps. Elle permet ainsi d’orienter les relations causales dans un graphe lorsqu’une cause est déjà connue. Si une variable prise à un temps $t - \tau$ cause une autre variable à un temps t , la relation causale est définie comme une relation causale décalée. Si la cause apparaît en même temps que son effet, la relation est dite instantanée.

Il existe plusieurs types de graphes pour représenter les relations causales entre séries temporelles comme le graphe causal fenêtre (*window causal graph*), Figure 1 (a) et le graphe causal résumé (*summary causal graph*), Figure 1 (b). Dans le premier type de graphe, un noeud représente une variable à un instant donné, tandis que le second ne donne que la variable : l'information temporelle est donc perdue. De telles représentations sont possibles grâce à la stationnarité causale [Runge, 2018] qui stipule que toutes les relations causales restent constantes en direction au cours du temps.

Un noeud X^r tel que $X^i \rightarrow X^r \leftarrow X^j$ est appelé un *collider* comme par exemple le noeud

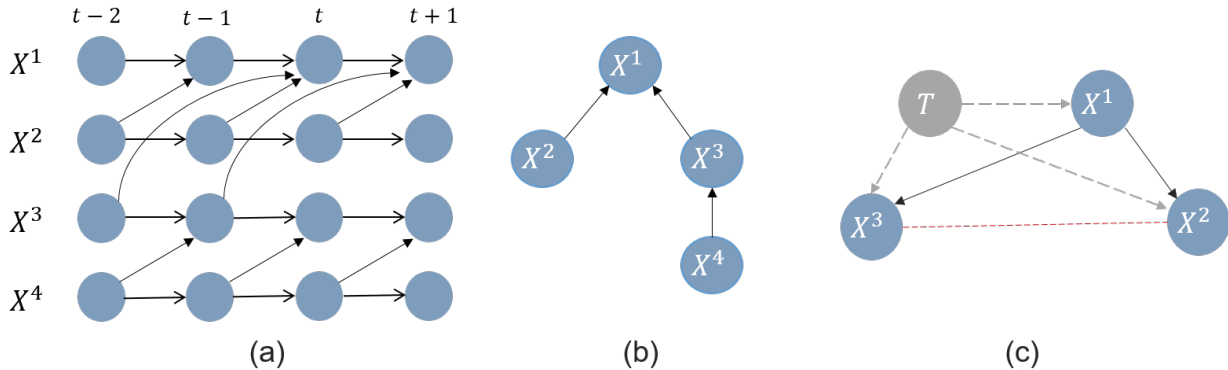


FIGURE 1 – Sur la gauche (a), un *window causal graph* et au milieu (b), le graphe causal *summary* correspondant. A droite (c), un autre *summary* où T est un facteur confondant non-observé.

X^1 sur la Figure 1 (b). Les antécédents d'un noeud dans le graphe sont appelés ses parents, X^1 a pour parents le couple (X^2, X^3) . Un noeud X^r appartient à l'ensemble de séparation de deux noeuds (X^i, X^j) si conditionner sur le premier rend le couple indépendant, par exemple le noeud X^3 d-sépare X^4 et X^1 dans la Figure 1 (b). Si une variable cause plusieurs noeuds, elle est appelée une cause commune (*confounder*), comme le noeud X^1 sur la Figure 1 (c). Une cause commune cachée est un *hidden confounder*, c'est le cas de la variable T sur cette même figure.

Pour constituer les arêtes du graphe, les critères d'indépendance ou d'indépendance conditionnelle sont primordiaux. Ils font en effet le lien entre des paramètres statistiques évalués dans les données et le graphe causal, par la propriété causale de Markov et la propriété de *faithfulness* [Peters et al., 2017]. Retrouver ces liens causaux est appelé la recherche de causalité (*causal discovery*). Il existe un grand nombre de méthodes de découverte causale catégorisées en différentes familles selon leur approche pour construire le graphe [Vowels et al., 2022] dont celle dite *constraint-based*. Cette dernière cherche à établir un graphe qui respecte au mieux les ensembles d'indépendance conditionnelle détectés dans les données.

2.2 Le problème des variables latentes

Dans la plupart des situations réelles, il est difficile d'affirmer qu'un jeu de données contient toutes les causes du système observé. Cela implique donc l'existence de variables latentes, ou non-observées. Ignorer ces variables peut conduire à des conclusions fausses et

des analyses incomplètes car elles peuvent introduire un biais significatif dans les résultats. Intégrer ces variables cachées dans les modèles permet donc de mieux étudier les mécanismes sous-jacents du système.

Pour cela, les DAG sont étendus pour contenir jusqu’à six types d’arêtes : une arête non-dirigées ($-$), une flèche seule (\rightarrow), une double flèche (\leftrightarrow), une arête non-dirigée d’un côté et indéterminée de l’autre ($\circ\rightarrow$) et enfin une arête indéterminée des deux côtés ($\circ-\circ$). Le graphe ainsi inféré est appelé un *Partial Ancestral Graph* (PAG). Dans un PAG, $X^1 \rightarrow X^2$ signifie que X^1 est une cause (potentiellement indirecte) de X^2 et que X^2 ne cause pas X^1 . Ainsi, une variable en cause une autre si elle est un antécédent de cette variable. Une double flèche entre deux noeuds indique qu’il n’y a pas de relation causale directe entre eux mais qu’une cause commune les influence. Munis de ces relations, certains algorithmes ont été établis sans l’hypothèse de suffisance causale. Parmi ceux-ci, FCI [Spirites et al., 2000, Zhang, 2008], un algorithme *constraint-based*, utilise les critères d’indépendance (conditionnelle) pour estimer un graphe causal et un ensemble de règles pour orienter ce graphe et détecter des variables latentes.

Une règle fondamentale, nommée *origin of causality*, permet d’orienter les triplets de variables $X^i \circ\leftarrow X^r \circ\rightarrow X^j$ tels que X^i et X^j ne sont pas adjacents, en *unshielded colliders*. Un *unshielded collider* est un noeud X^r tel que $X^i \circ\rightarrow X^r \leftarrow X^j$, si X^r n’est pas dans l’ensemble de séparation de X^i et X^j . Cette règle, souvent mentionnée comme la règle R_0 , permet de former des ensembles de *Possible D-separation* [Spirites et al., 2000] entre deux noeuds. Ces ensembles, présents dans les PAG, se comportent de manière similaire aux parents de noeuds dans un DAG.

2.3 Séries temporelles

Les séries temporelles sont des données ordonnées qui présentent souvent des structures d’auto-corrélation. Cela signifie que leur passé influence leur présent et leur futur. Il est nécessaire de les prendre en compte lors de la construction du graphe causal. Le décalage temporel entre deux observations permet d’orienter le graphe : une cause ne peut intervenir avant son effet. Le temps peut être vu comme un facteur confondant entre plusieurs variables : si chaque variable entretient une relation linéaire avec le temps alors elles partagent un point commun. Cela peut donc créer un lien entre ces variables, comme observé sur la Figure 1 (c). Sur cette dernière, le temps est représenté par la variable T et agit sur les séries X^1 , X^2 et X^3 . Si T n’est pas considérée dans le modèle, cela peut induire une corrélation entre ces variables et créer un lien entre X^2 et X^3 qui ne devrait pas exister.

Plusieurs algorithmes ont été proposés pour construire un graphe causal à partir de séries temporelles comme VarLiNGAM [Hyvärinen et al., 2010], PCMCi [Runge et al., 2019b], PCGCE [Assaad et al., 2022a] ou encore PC-PMIME [Arsac et al., 2023]. Plusieurs de ces méthodes s’appuient sur l’indépendance conditionnelle.

2.4 Tester l'indépendance conditionnelle

Deux variables aléatoires X et Y sont conditionnellement indépendantes à une troisième variable Z si et seulement si (ssi) $P_{X,Y|Z} = P_{X|Z}P_{Y|Z}$, on note $X \perp\!\!\!\perp Y \mid Z$. Intuitivement, cela signifie que la distribution de X sachant Y et Z est en fait complètement déterminée par la valeur de Z seule, Y devient superflue lorsque Z est donnée.

Plusieurs tests d'indépendance conditionnelle sont basés sur la théorie de l'information qui permet d'établir des tests non-paramétriques et donc d'être appliquée à de nombreux modèles de données. Pour les séries chronologiques, des mesures ont été développées à partir de l'Information Mutuelle Conditionnelle (CMI) comme l'entropie de transfert partielle (PTE) [Vakorin et al., 2009] ou la mesure PMIME [Kugiumtzis, 2013].

La mesure PMIME est asymétrique, non-paramétrique et conçue pour détecter des couplages directs au sein de séries temporelles. Elle est dérivée d'un processus d'*embedding* basé sur un critère de sélection, la CMI. Dans le cas multivarié, pour évaluer si une variable X a une influence sur une variable Y , conditionnellement à un ensemble de variables $\mathbf{Z} = (Z^1, \dots, Z^{g-2})$, le processus construit itérativement un vecteur d'*embedding* \mathbf{w} à partir des composantes décalées extraites de (X, Y, \mathbf{Z}) qui expliquent le mieux le futur de Y , noté $Y_t^T = (Y_{t+1}, \dots, Y_{t+T})$. Plus précisément, le vecteur \mathbf{w} est construit en piochant successivement des vecteurs dans la matrice $W = (X_{t-1}, \dots, X_{t-\tau_{max}}, Y_{t-1}, \dots, Y_{t-\tau_{max}}, Z_{t-1}^1, \dots, Z_{t-\tau_{max}}^{g-2})$. Chaque itération définit un cycle d'*embedding* et utilise un critère d'arrêt, noté A , pour accepter ou refuser une composante. Une composante est acceptée si l'information qu'elle apporte améliore strictement l'information déjà contenue dans le vecteur d'*embedding*.

Ainsi, le vecteur \mathbf{w} est formé de g' variables décalées dans le temps sélectionnées par la CMI et peut se décomposer en $\mathbf{w}_t = (w_t^x, w_t^y, w_t^z)$, où w_t^x sont les composantes de X sélectionnées dans le procédé, w_t^y , celles de Y et les composantes restantes sont regroupées sous w_t^z . On quantifie alors l'effet causal de X vers Y conditionnellement à \mathbf{Z} par un ratio entre une CMI et une information mutuelle :

$$R_{X \rightarrow Y | \mathbf{Z}} = \frac{I(Y_t^T, \mathbf{w}_t^x | \mathbf{w}_t^y, \mathbf{w}_t^z)}{I(Y_t^T; \mathbf{w}_t)}.$$

La description détaillée de la construction du vecteur \mathbf{w} est donnée dans [Kugiumtzis, 2013]. Si w_t^x est vide cela signifie que X n'a aucune influence sur Y , ce qui se transcrit sur la mesure R : si w_t^x est vide, alors R est nulle. De plus, la mesure est bornée entre 0 et 1, 0 signifie indépendance et 1 signifie que le futur de Y est totalement déterminé par X .

3 Approche proposée

3.1 Motivations et hypothèses

Dans ce travail nous considérons un système modélisé par des séries temporelles caractérisé par des relations linéaires ou non-linéaires entre les variables, sans faire d'hypothèse spécifique sur leur distribution *a priori*. Nous supposons l'existence de causes cachées communes (*hidden confounders*), écartant ainsi l'hypothèse de suffisance causale. Le modèle des séries tempo-

relles considéré est général, décrivant une variable X^i observée au temps t selon :

$$X_t^i = f(X_{t-\gamma_{1,1}}^{p_1}, X_{t-\gamma_{1,2}}^{p_1}, \dots, X_{t-\gamma_{1,j_1}}^{p_1}, X_{t-\gamma_{2,1}}^{p_2}, \dots, X_{t-\gamma_{k,j_k}}^{p_k}, \varepsilon_t^i),$$

où p_1, \dots, p_k sont les indices des variables parentes de X^i et $\gamma_{l,m}$ représente le décalage m avec lequel la variable X^{p_l} influence X_t^i , tel que $\gamma_{l,1} > \gamma_{l,2} > \dots \geq 0$. Les $(\varepsilon_t^i)^n$, indépendants les uns des autres et indépendants des causes de X^i , représentent un bruit ajouté.

Pour construire un graphe causal à partir de séries temporelles dans ce cadre, différentes méthodes existent comme tsFCI [Entner and Hoyer, 2010], une extension de FCI capable de traiter des séries temporelles. Cependant, ses auteurs indiquent que tsFCI n'est pas fiable pour pratiquement toutes les tailles de séries temporelles. D'autres extensions existent également comme SVAR-GFCI [Malinsky and Spirtes, 2018], qui se concentre sur les modèles auto-régressifs structurels linéaires ou LPCMCI [Gerhardus and Runge, 2020]. Ce dernier modèle semble respecter nos conditions de travail mais souffre d'un temps de calcul conséquent lors de son utilisation avec une mesure de dépendance non-linéaire, la CMI, et offre une faible interprétabilité.

Nous proposons donc d'utiliser un algorithme de type FCI avec la mesure d'indépendance conditionnelle PMIME. La méthode, nommée FCI-PMIME, permet de construire un graphe causal à partir de données de séries temporelles en l'absence de suffisance causale. Cet algorithme traite des relations linéaires et non-linéaires entre les variables. De plus, [Papana et al., 2013] montre que la mesure PMIME détecte mieux que la PTE, les dépendances et indépendances conditionnelles dans le cas de systèmes non-linéaires, ce qui s'avère crucial lors de l'utilisation d'une méthode *constraint-based*. Un avantage certain de la mesure PMIME est qu'elle prend en compte directement l'aspect temporel des variables, en considérant une fenêtre temporelle lors du test, contrairement à la CMI par exemple. Enfin, la mesure est bornée entre 0 et 1, simplifiant son interprétabilité et évitant l'ajout d'un test de signification statistique supplémentaire.

3.2 L'algorithme FCI-PMIME

L'algorithme FCI-PMIME fonctionne de façon analogue aux méthodes *constraint-based*. Il essaye dans un premier temps de trouver les ensembles d'indépendance et d'indépendance conditionnelle dans les données puis applique des règles d'orientation spécifiques pour établir le graphe causal. Une caractéristique essentielle de FCI-PMIME est sa capacité à mesurer des interactions décalées dans le temps ainsi que des relations instantanées. La mesure PMIME est utilisée pour évaluer les dépendances du passé vers le futur, tandis que pour les relations instantanées, l'algorithme recourt à la MI ou la CMI [Runge, 2018].

L'algorithme commence par former un graphe complet $G = (V, E)$, où $V = (V_{t-}, V_t)$ représente les noeuds du passé et du présent et E , les arêtes. Une phase d'élagage du graphe débute dans laquelle une arête entre X et Y est enlevée si $R_{X \rightarrow Y} = 0$, où R est la mesure PMIME ou la CMI. Lorsque toutes les arêtes ont été testées et que certaines ont été retirées, les tests se poursuivent pour celles qui restent en regardant si deux noeuds sont conditionnellement indépendants. L'ensemble de conditionnement se compose dans un premier temps d'une seule variable parmi les parents de X ou Y , puis sa taille augmente incrémentalement jusqu'à ce

qu'une indépendance conditionnelle soit détectée ou que toutes les arêtes liées à X et Y ont été testées. Comme PMIME est asymétrique, l'algorithme teste les deux directions : de X vers Y et de Y vers X .

A l'issue de cette première phase, le graphe obtenu est non-dirigé. La règle R_0 présentée en section 2.2 est alors mise en oeuvre pour obtenir une pré-orientation du graphe et identifier les *unshielded colliders*. Dans une troisième phase, les ensembles de *Possible-d-sep* sont établis. La première phase est alors réitérée en conditionnant cette fois par les variables se trouvant dans les ensembles de *Possible-d-sep* de X et Y plutôt que par les variables adjacentes. Enfin, l'algorithme termine par l'utilisation d'un ensemble de règles qui permettent d'orienter

Algorithme 1 : FCI-PMIME

Entrées : n observations de X^0, \dots, X^g , paramètres de PMIME

Sorties : \mathcal{G} , le graphe estimé

Créer un graphe complet non-dirigé $\mathcal{G} = (V, E)$, avec $V = (V_{t-}, V_t)$ et tel que

$$X_{t-}^i \circ \rightarrow X_t^j, \quad \forall X_{t-}^i \in V_{t-}, X_t^j \in V_t$$

pour chaque permutation (X^i, X^j) de noeuds \mathcal{G} **faire**

└ Calculer la mesure R entre X^j et X^i , affecter R à l'arête $V^{i,j}$

Retirer les arêtes $V^{i,j}$ de V si $R \approx 0$

Initialiser $l = 1$, $process = \text{Vrai}$, $sep_set = []$

tant que $process$ est *Vrai* **faire**

└ Définir $process$ sur Faux

└ **pour** chaque permutation $(X^i, X^j) \in V$ **faire**

└└ $adj_set = Par(X^i, \mathcal{G}) \setminus X^j$ la liste des prédécesseurs de X^i , sans X^j

└└ **si** $Card(adj_set) \geq l$ **alors**

└└└ **pour** chaque combinaison $\mathbf{Z} \in adj_set$ de taille l **faire**

└└└└ Calculer $R(X^j \rightarrow X^i \mid \mathbf{Z})$, affecter R à l'arête entre X^j et X^i

└└└└ **si** $R \approx 0$ **alors**

└└└└└ ajouter \mathbf{Z} à sep_set

└└└ Définir $process$ comme *Vrai*

Retirer les arêtes $V^{i,j}$ de V si $R \approx 0$

Appliquer la règle d'orientation R_0 pour détecter les *unshielded colliders*

Identifier les ensembles de *Possible-D-separation*

Répéter la boucle Tant que précédente en itérant sur les membres de *Possible D-separation* entre X^i et X^j

Réinitialiser l'orientation du graphe $\mathcal{G} : X^i \circ \rightarrow X^j, \forall X^i, X^j \in V$

Orienter \mathcal{G} en appliquant la règle R_0 puis les règles 1 à 4 et 8 à 10 de FCI.

le graphe. L'orientation du graphe est réinitialisée et la règle R_0 est de nouveau appliquée. Ensuite, les règles 1, 2, 3 et 4, de FCI [Spirtes et al., 2000] sont employées ainsi que les règles 8, 9, 10 de [Zhang, 2008]. Les règles 5 à 7 de ce dernier traitent des effets cachés et ne sont pas prises en compte dans notre contexte car nous ne nous intéressons qu'aux causes communes cachées. Certaines règles sont aussi adaptées pour les séries temporelles : par exemple, toutes les arêtes telles que $X_{t-\tau}^i \circ \rightarrow X_t^j$ sont orientées $X_{t-\tau}^i \rightarrow X_t^j$, indiquant simplement que la causalité ne peut se propager dans le passé. Notre méthode, FCI-PMIME est décrit dans l'Algorithme 1.

La métrique employée pour comparer les différentes méthodes est le $F1$ -score, où un vrai-positif apparaît lorsqu’une arête dans le graphe estimé est aussi dans le graphe simulé. L’auto-corrélation n’est pas prise en compte dans la mesure du score car toutes les méthodes parviennent à la détecter ce qui fausse grandement le score. Sont alors calculées les moyennes des $F1$ -scores obtenus sur 10 graphes simulés pour différentes longueurs de séries temporelles $n = \{125, 250, 500, 1000, 2000, 4000\}$.

Pour chaque longueur de série, la méthode FCI-PMIME obtient des performances supérieures aux autres méthodes. De plus, son score augmente avec la taille des séries temporelles, contrairement à l’algorithme PC-PMIME, par exemple, ce qui montre une certaine consistance. Une légère instabilité est constatée pour des séries de faible taille, expliquée par l’estimation asymptotique de la CMI dans la mesure PMIME. Ainsi, lorsque le nombre d’observation est réduit, les indépendances (conditionnelles) sont plus difficiles à détecter et le graphe inféré comporte souvent un nombre d’arête sur-estimé.

5 Conclusion et futurs travaux

Ce papier présente une méthode pour apprendre un graphe causal à partir de données de séries temporelles dans un système complexe où toutes les variables ne sont pas observées. Notre approche, FCI-PMIME, limite le nombre d’hypothèses sur les données et permet d’étudier des relations entre linéaires ou non-linéaires entre variables et décalées dans le temps. La méthode produit des résultats prometteurs sur des données simulées, qui indiquent toutefois qu’une marge de progression est possible. Un test sur des données réelles permettrait d’évaluer la méthode sur des données plus complexes. Dans le domaine de la fabrication, l’objectif consisterait, par exemple, à identifier les facteurs responsables d’une défaillance au sein d’un système complexe sous surveillance.

La plupart des méthodes de découverte causale existantes souffre d’un problème de passage à l’échelle, et la nôtre n’y fait pas exception. En effet, les méthodes peuvent se comporter correctement sur des systèmes composés de peu de variables mais lorsque ce nombre devient important, un manque de stabilité s’observe ainsi qu’une diminution de l’interprétabilité du graphe estimé. Une solution serait par exemple de traiter les dépendances entre groupes de variables plutôt que variable par variable.

Enfin, il s’avère que les tests d’indépendance conditionnelle dans un cadre général souffrent d’une faible puissance statistique [Shah and Peters, 2020], c’est aussi le cas pour la mesure PMIME. Cependant, nous pensons que le critère de sélection et le processus *d’embedding* réalisé dans la méthode permettent d’augmenter cette puissance statistique. De prochains travaux pourraient être d’obtenir des garanties théoriques sur la puissance du test basé sur ce critère de sélection.

Références

[Arsac et al., 2023] Arsac, A., Lomet, A., and Poli, J.-P. (2023). Causal discovery for time series with constraint-based model and pmime measure. In *When Causal Inference meets Statistical Analysis*.

- [Assaad et al., 2022a] Assaad, C. K., Devijver, E., and Gaussier, E. (2022a). Inferring extended summary causal graphs from observational time series. *arXiv preprint arXiv :2205.09422*.
- [Assaad et al., 2022b] Assaad, C. K., Devijver, E., and Gaussier, E. (2022b). Survey and evaluation of causal discovery methods for time series. *Journal of Artificial Intelligence Research*, 73 :767–819.
- [Brodersen et al., 2015] Brodersen, K. H., Gallusser, F., Koehler, J., Remy, N., and Scott, S. L. (2015). Inferring causal impact using bayesian structural time-series models. *The Annals of Applied Statistics*, 9(1) :247–274.
- [Entner and Hoyer, 2010] Entner, D. and Hoyer, P. O. (2010). On causal discovery from time series data using fci. *Probabilistic graphical models*, pages 121–128.
- [Frenzel and Pompe, 2007] Frenzel, S. and Pompe, B. (2007). Partial mutual information for coupling analysis of multivariate time series. *Physical review letters*, 99(20) :204101.
- [Gerhardus and Runge, 2020] Gerhardus, A. and Runge, J. (2020). High-recall causal discovery for auto-correlated time series with latent confounders. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M., and Lin, H., editors, *Advances in Neural Information Processing Systems*, volume 33, pages 12615–12625. Curran Associates, Inc.
- [Hyvärinen et al., 2010] Hyvärinen, A., Zhang, K., Shimizu, S., and Hoyer, P. O. (2010). Estimation of a structural vector autoregression model using non-gaussianity. *Journal of Machine Learning Research*, 11(5).
- [Kugiumtzis, 2013] Kugiumtzis, D. (2013). Direct-coupling information measure from nonuniform embedding. *Physical Review E*, 87(6) :062918.
- [Malinsky and Spirtes, 2018] Malinsky, D. and Spirtes, P. (2018). Causal structure learning from multivariate time series in settings with unmeasured confounding. In *Proceedings of 2018 ACM SIGKDD workshop on causal discovery*, pages 23–47. PMLR.
- [Pamfil et al., 2020] Pamfil, R., Sriwattanaworachai, N., Desai, S., Pilgerstorfer, P., Georgatzis, K., Beaumont, P., and Aragam, B. (2020). Dynotears : Structure learning from time-series data. In *International Conference on Artificial Intelligence and Statistics*, pages 1595–1605. PMLR.
- [Papana et al., 2013] Papana, A., Kyrtsov, C., Kugiumtzis, D., and Diks, C. (2013). Simulation study of direct causality measures in multivariate time series. *Entropy*, 15(7) :2635–2661.
- [Pearl, 2009] Pearl, J. (2009). Causal inference in statistics : An overview. *Statistics surveys*, 3 :96–146.
- [Peters et al., 2017] Peters, J., Janzing, D., and Schölkopf, B. (2017). *Elements of causal inference : foundations and learning algorithms*. The MIT Press.
- [Runge, 2018] Runge, J. (2018). Conditional independence testing based on a nearest-neighbor estimator of conditional mutual information. In *International Conference on Artificial Intelligence and Statistics*, pages 938–947. PMLR.
- [Runge et al., 2019a] Runge, J., Bathiany, S., Bollt, E., Camps-Valls, G., Coumou, D., Deyle, E., Glymour, C., Kretschmer, M., Mahecha, M. D., Muñoz-Marí, J., et al. (2019a). Inferring causation from time series in earth system sciences. *Nature communications*, 10(1) :2553.
- [Runge et al., 2019b] Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., and Sejdinovic, D. (2019b). Detecting and quantifying causal associations in large nonlinear time series datasets. *Science advances*, 5(11) :eaau4996.
- [Shah and Peters, 2020] Shah, R. D. and Peters, J. (2020). The hardness of conditional independence testing and the generalised covariance measure.
- [Spirtes et al., 2000] Spirtes, P., Glymour, C. N., Scheines, R., and Heckerman, D. (2000). *Causation, prediction, and search*. MIT press.
- [Vakorin et al., 2009] Vakorin, V. A., Krakovska, O. A., and McIntosh, A. R. (2009). Confounding effects of indirect connections on causality estimation. *Journal of neuroscience methods*, 184(1) :152–160.
- [Vowels et al., 2022] Vowels, M. J., Camgoz, N. C., and Bowden, R. (2022). D’ya like dags? a survey on structure learning and causal discovery. *ACM Computing Surveys*, 55(4) :1–36.
- [Zhang, 2008] Zhang, J. (2008). On the completeness of orientation rules for causal discovery in the presence of latent confounders and selection bias. *Artificial Intelligence*, 172(16-17) :1873–1896.