

**Wydział Elektroniki i Technik Informacyjnych
Politechnika Warszawska**

Seminarium dyplomowe magisterskie 1

Przegląd literatury

inż. Hubert Pisiecki

Warszawa, 2024

Spis treści

1	Wstęp	5
2	Analiza przyczynowości - przegląd teoretyczny	7
2.1	Definicja i rola analizy przyczynowości	7
3	Klasyczne metody analizy przyczynowości	9
3.1	Korelacja krzyżowa (Cross corelation)	9
3.2	Przyczynowość w sensie Grangera (Granger Causality)	10
3.3	Częściowa koherencja ukierunkowana (Partial Directed Coherence)	11
4	Transfer Entropy	13
	Bibliografia	15

Rozdział 1

Wstęp

Współczesny przemysł charakteryzuje się wysokim stopniem złożoności, wynikającym z licznych powiązań i interakcji pomiędzy różnymi komponentami systemów na dużą skalę. W takich środowiskach nawet drobna usterka może rozprzestrzeniać się wzdłuż ścieżek przepływu informacji i materiałów, wpływając na inne elementy systemu, a w konsekwencji powodując poważne zakłócenia w całym procesie. Wpływ takich zakłóceń może prowadzić do obniżenia wydajności, zwiększenia kosztów operacyjnych, a w najgorszych przypadkach do nieplanowanych przestojów zakładów produkcyjnych lub nawet ich uszkodzeń.

Kluczowym wyzwaniem współczesnego przemysłu jest szybka i precyzyjna detekcja oraz diagnozowanie zakłóceń w procesach technologicznych. Wraz z rosnącą automatyzacją i cyfryzacją, liczba alarmów generowanych w systemach produkcyjnych dynamicznie wzrasta, co sprawia, że zarządzanie nimi staje się coraz bardziej skomplikowane. Tradycyjne metody diagnostyki, choć sprawdzone w wielu zastosowaniach, często nie nadążają za złożonością współczesnych procesów, zwłaszcza gdy dane charakteryzują się nieliniowością, niestacjonarnością i obecnością zakłóceń.

Awarie mogą propagować się na różne części systemu, tworząc kaskady alarmów, które utrudniają operatorom zlokalizowanie źródła problemu. W efekcie podjęcie właściwej decyzji w krótkim czasie staje się dużym wyzwaniem. Dlatego coraz większe znaczenie zyskują nowoczesne metody analizy danych procesowych, które nie tylko pozwalają na bieżącą obserwację stanu systemu, ale również na identyfikację ukrytych zależności przyczynowo-skutkowych między różnymi komponentami (Yu and Yang, 2015).

Jednym z obiecujących podejść jest analiza przyczynowości, która umożliwia śledzenie, jak nieprawidłowości w jednym obszarze procesu wpływają na inne części systemu. Narzędzia takie jak Transfer Entropy pozwalają na wykrywanie siły i kierunku przepływu informacji między zmiennymi, co może znacząco usprawnić wykrywanie anomalii i ograniczyć liczbę fałszywych alarmów (Lindner et al., 2017a).

Integracja zaawansowanych metod analitycznych z istniejącymi systemami alarmowymi otwiera drogę do budowy inteligentnych systemów ostrzegania, które potrafią nie tylko reagować na bieżące problemy, ale także przewidywać potencjalne zagrożenia, zanim wpłyną one na cały proces.

W przemyśle chemicznym analiza przyczynowości umożliwia identyfikację źródeł zakłóceń w procesach, co pozwala na ich efektywne eliminowanie i poprawę jakości produktów. Przykładowo, badania Lindnera, Auret i Bauer (Lindner et al., 2017b) analizują wpływ perturbacji w procesach chemicznych na techniki analizy przyczynowości, podkreślając znaczenie TE w takich analizach.

Transfer Entropy mierzy przepływ informacji między zmiennymi i pozwala na wychwycenie nieliniowych zależności, co wyróżnia je na tle tradycyjnych metod, takich jak korelacja krzyżowa czy przyczynowość w sensie Grangera. Badania Yu i Yanga (Yu and Yang, 2015) pokazują zastosowanie TE do wykrywania przyczynowości między zmiennymi procesowymi na podstawie danych alarmowych w przemyśle.

Rozdział 2

Analiza przyczynowości - przegląd teoretyczny

Analiza przyczynowości to obszar badań w statystyce i projektowaniu eksperymentów, który ma na celu ustalenie związku przyczynowo-skutkowego. Zwykle obejmuje ona cztery kluczowe elementy: korelację, kolejność w czasie (przyczyna musi wystąpić przed skutkiem), wiarygodny mechanizm fizyczny lub teoretyczny, który wyjaśnia dlaczego dany skutek wynika z określonej przyczyny oraz eliminację innych możliwych przyczyn, które mogą wyjaśniać zaobserwowany efekt. Badanie i wyjaśnianie związków przyczynowo-skutkowych pomiędzy zmiennymi, zdarzeniami lub obiektami stanowiło kluczowe zagadnienie w wielu dziedzinach naukowych, między innymi w ekonomii (Imbens, 2004) czy naukach medycznych (Mani and Cooper, 2000).

2.1 Definicja i rola analizy przyczynowości

Nie ma jednej powszechnie uznawanej definicji przyczynowości. Można ją rozumieć jako "przepływ" między procesami, wyrażony matematycznie i poddany analizie. Wczesniejsze badania koncentrowały się na przyczynach w sytuacjach deterministycznych, gdzie kluczowe są dwa warunki: konieczność – jeśli X występuje, to Y musi wystąpić, oraz wystarczalność – jeśli Y występuje, to X musiał wystąpić. Jednak deterministyczne podejście nie odzwierciedla rzeczywistości, ponieważ żaden system nie jest całkowicie deterministyczny. Dlatego bardziej realistyczne jest podejście probabilistyczne, które mówi, że jeśli X występuje, to prawdopodobieństwo Y rośnie. Pierwsza definicja przyczynowości, zaproponowana przez Wienera w 1956 roku brzmi: „X można uznać za 'przyczynę' Y, jeśli przewidywalność Y poprawia się po uwzględnieniu informacji o X” (Wiener, 1956). Koncepcja ta, zastosowana do analiz szeregów czasowych, została rozwinięta przez Clive'a W. J. Grangera (Granger, 2004), laureata Nagrody Nobla w 2003 roku.

W fizyce i nieliniowej dynamice rośnie zainteresowanie badaniem współzależności między sprzężonymi, złożonymi systemami (Boccaletti et al., 2002). Nieliniowe rozszerzenie przyczynowości

Grangera, wykorzystujące lokalne predyktory liniowe, zaproponowali Chen i in. w 2004 roku (Chen et al., 2004). Ważną klasą nieliniowych predyktorów są funkcje bazowe promieniowe (Butte and Kohane, 2000), używane do nieliniowego rozszerzenia koncepcji przyczynowości Grangera (Ancona et al., 2004). Thomas Schreiber w 2000 roku w swojej pracy poświęconej nieliniowej analizie dynamiki systemów (Schreiber, 2000), zaproponował nieparametryczną metodę do pomiaru transferu informacji przyczynowej między systemami – Transfer Entropy (TE), która jest miarą opartą na odległości Kullbacka-Leiblera (mierzy asymetryczną różnicę między dwoma rozkładami prawdopodobieństwa). TE jest stosowana między innymi w neurologii i neuroinformatyce (Lizier et al., 2011), sieciach społecznościowych (Stegg and Galstyan, 2012) czy badaniu konfliktów zbrojnych (Kushwaha and Lee, 2023).

Połączenia przyczynowe między parą zmiennych można wykryć analizując relacje między ich szeregami czasowymi. Na tej podstawie można wygenerować graf przyczynowy, gdzie węzły reprezentują zmierzone zmienne, a krawędzie między nimi odzwierciedlają wpływ przyczynowy między tymi zmiennymi. Podejście to ma jednak pewne ograniczenia. Po pierwsze, awarie nie zawsze są bezpośrednio związane z jedną mierzoną zmienną. Analiza przyczynowa z wykorzystaniem tego grafu może wskazać zmienną, sekcję lub pętlę regulacji, która jest silnie powiązana z awarią, umożliwiając dalszą analizę w celu zidentyfikowania jej źródła. Awarye mogą również zmieniać strukturę przyczynową procesu. Na przykład awaria zaworu w pętli regulacyjnej może zniekształcić połączenie przyczynowe między zmienną sterowaną a manipulowaną (Lindner et al., 2017a). Analiza przyczynowa za pomocą grafu przyczynowego pozwala więc wykrywać powiązania między zmiennymi, ale jej skuteczność jest ograniczona, gdy awarye zmieniają strukturę procesu lub nie są związane z jedną zmienną.

Techniki analizy danych procesowych umożliwiają identyfikację powiązań przyczynowych między zmiennymi na podstawie ich szeregów czasowych, a metody te można podzielić na trzy główne kategorie: metody opóźnieniowe, metody oparte na niezależności warunkowej oraz metody wykorzystujące wyższe statystyki. Metody oparte na opóźnieniach uwzględniają informacje tymczasowe poprzez uwzględnienie w analizie wcześniejszych pomiarów. Sprawdzają się one dobrze przy analizie związku przyczynowego pomiędzy szeregami czasowymi i dlatego są najszerszej stosowane w monitorowaniu procesów (Yang et al., 2014).

Rozdział 3

Klasyczne metody analizy przyczynowości

W literaturze występuje kilka klasycznych metod badania analizy przyczynowości. Są to między innymi Korelacja krzyżowa (CC), przyczynowość w sensie Grangera (GC), a także częściowa koherencja ukierunkowana (PDC) (Lindner et al., 2017a).

3.1 Korelacja krzyżowa (Cross corelation)

Dostosowuje szeregi czasowe X i Y , aby występowało między nimi opóźnienie. Współczynnik korelacji pomiędzy skorygowanymi szeregami czasowymi można obliczyć dla zakresu zakładanego opóźnienia. Korelacja ρ dla założonego opóźnienia k jest obliczana w następujący sposób:

$$\rho_k = \frac{1}{N-k} \sum_{i=1}^{N-k} \frac{(x_i - \mu_x)(y_{i+k} - \mu_y)}{\sigma_x \sigma_y} \quad (1)$$

Gdzie:

N - liczba próbek

μ - średnia szeregu czasowego

σ - odchylenie standardowe szeregu czasowego

Zakłada się, że maksymalna znaleziona korelacja jest rzeczywistą korelacją pomiędzy szeregami czasowymi, a opóźnienie które daje tę maksymalną wartość korelacji, przyjmuje się jako oszacowanie opóźnienia czasowego między zmiennymi. Zaletą korelacji krzyżowej jest przede wszystkim jej prostota i niewymagające obliczenia. Jednakże, korelacja krzyżowa może wykryć jedynie liniową interakcję pomiędzy szeregami czasowymi. W procesach, w których występuje zachowanie nieliniowe, wyniki

mogą być niedokładne (Bauer and Thornhill, 2008). Trend w szeregu czasowym jest ignorowany w analizie korelacji krzyżowej; wartości w różnych momentach czasowych traktowane jako próbki tego samego zdarzenia losowego (Yang et al., 2014). Innymi słowy, gdy korelacja między X i Y jest obliczana przy określonym opóźnieniu k , wartości szeregów czasowych X i Y uważa się za powtarzające się pomiary tego samego zdarzenia, zamiast postrzegać je jako pomiary różnych zdarzeń w czasie (Lindner et al., 2017a).

3.2 Przyczynowość w sensie Grangera (Granger Causality)

Clive Granger opracował formalizację opisu przyczynowości Wienera w kontekście modeli regresji, zwane „przyczynowością Grangera” (Granger, 2004). Z założenia mówi o tym, że X ma przyczynowy wpływ na Y , gdy włączenie przeszłych pomiarów X do modelu regresji Y skutkuje poprawą przewidywań autoregresyjnego modelu Y (Wiener, 1956). Rozważając szeregi czasowe dwóch zmiennych $X(t)$ i $Y(t)$, $Y(t)$ może być zamodelowane jako model autoregresyjny (AR), określane jako model ograniczony:

$$Y(t) = \sum_{j=1}^k B_j Y(t-j) + \varepsilon(t) \quad (2)$$

W tym przypadku uwzględniane są tylko przeszłe wartości Y , aby przewidzieć przyszłe wartości tego samego Y . W Równaniu 2 k oznacza rząd modelu, który definiuje opóźnienie czasowe; B jest współczynnikiem AR, a ε_y to błąd predykcji. $Y(t)$ może być również modelowane z uwzględnieniem przeszłych wartości $X(t)$ i $Y(t)$, co jest znane jako model pełny, jak pokazano w Równaniu 3:

$$Y(t) = \sum_{j=1}^k [A_{yx,j} X(t-j) + A_{yy,j} Y(t-j)] + \varepsilon_{y|x}(t) \quad (3)$$

Gdzie A_{yx} i A_{yy} reprezentują współczynniki wielomianu modelu. A_{yx} , A_{yy} oraz B można określić za pomocą metody najmniejszych kwadratów. Rząd modelu k można ustalić za pomocą Kryterium Informacyjnego Akaikiego. Gdy wariancja $\varepsilon_{y|x}(t)$ jest mniejsza niż wariancja $\varepsilon_y(t)$, oznacza to, że przewidywane Y jest ulepszone przez uwzględnienie wcześniejszych wartości X . Mówi się wówczas, że X podowuje Y w sensie Grangera. Przyczynowość Grangera można określić ilościowo, jak pokazano w Równaniu 4:

$$F_{x \rightarrow y} = \ln \frac{\text{var}(\varepsilon_y)}{\text{var}(\varepsilon_{y|x})} \quad (4)$$

Gdy przewidywane Y **nie jest ulepszone** przez uwzględnienie X , $\text{var}(\varepsilon_{y|x}) = \text{var}(\varepsilon_y)$ i $F_{x \rightarrow y} = 0$. Gdy przewidywane Y **jest ulepszone** przez uwzględnienie X , $\text{var}(\varepsilon_{y|x}) < \text{var}(\varepsilon_y)$ i $F_{x \rightarrow y} > 0$. Przyczynowość Grangera dostarcza miary przyczynowości, której obliczenia są proste i niewymagające obliczeniowo. Koncepcja, z której pochodzi ta miara jest łatwa do zrozumienia poprzez intuicję,

ponieważ regresja jest dobrze znana. Wadą przyczynowości Grangera jest to, że miara przyczynowości zależy od dokładności modeli AR. Przyczynowość Grangera opiera się na liniowych modelach interakcji szeregów czasowych, które najpierw muszą zostać dopasowane do danych, dlatego nieliniowe interakcje w procesach mogą powodować niedokładność tej metody (Lindner et al., 2017a).

3.3 Częściowa koherencja ukierunkowana (Partial Directed Coherence)

Metody analizy przyczynowości w dziedzinie częstotliwości przedstawiają transfer energii między parami szeregów czasowych dla każdej częstotliwości (Landman et al., 2014). PDC została opracowana w celu zapewnienia opisu przyczynowości Grangera w dziedzinie częstotliwości (Baccalá and Sameshima, 2001). Szeregi czasowe mogą być modelowane za pomocą N-wymiarowego modelu autoregresyjnego, jak pokazano w Równaniu 5:

$$\begin{bmatrix} x_1(t) \\ \vdots \\ x_N(t) \end{bmatrix} = \sum_{r=1}^k A_r \begin{bmatrix} x_1(t-r) \\ \vdots \\ x_N(t-r) \end{bmatrix} + \begin{bmatrix} \epsilon_1(t) \\ \vdots \\ \epsilon_N(t) \end{bmatrix} \quad (5)$$

Odpowiedź częstotliwościowa procesu może być uzyskana poprzez zastosowanie transformaty Z (dyskretnej wersji transformaty Laplace'a) i ustawienie $z^{-1} = e^{-j\omega}$, co prowadzi do Równania 6:

$$A(\omega)X(\omega) = E(\omega) \quad (6)$$

$$A_{ij}(\omega) = - \sum_{r=1}^k a_{ij}(r) e^{-j\omega r} \quad (7)$$

$$X(\omega) = x_1(\omega) x_2(\omega) \dots x_N(\omega) \quad (8)$$

$$E(\omega) = \epsilon_1(\omega) \epsilon_2(\omega) \dots \epsilon_N(\omega) \quad (9)$$

Gdy $a_{ij}(r) = 0$ dla wszystkich wartości r , istnieją statystyczne dowody na brak przyczynowości od x_i do x_j . Opierając się na tej koncepcji, PDC jest zdefiniowana przez Równanie 10:

$$|\hat{\pi}_{if}(\omega)| = \frac{|A_{ij}(\omega)|}{\sqrt{\sum_{i=1}^n |A_{ij}(\omega)|^2}} \quad (10)$$

Ponieważ PDC opiera się na obliczeniach w dziedzinie częstotliwości, może być szczególnie przydatna w zastosowaniach, w których proces wykazuje zachowanie oscylacyjne (Lindner et al., 2017a).

Rozdział 4

Transfer Entropy

Transfer Entropy jest interpretacją definicji przyczynowości Wienera (Wiener, 1956) w ujęciu teorii informacji. TE z X do Y można obliczyć, jak pokazano w Równaniu 11:

$$T_{x \rightarrow y} = \sum p(y_{i+h}, Y_i^{(K)}, X_i^{(L)}) \cdot \log \frac{p(y_{i+h}|Y_i^{(K)}, X_i^{(L)})}{p(y_{i+h}|Y_i^{(K)})} \quad (11)$$

Gdzie $p(a, b)$ oznacza funkcję wspólnego rozkładu prawdopodobieństwa między a i b ; $p(a|b)$ oznacza warunkową funkcję rozkładu prawdopodobieństwa dla a przy danym b , a h to horyzont predykcji; $X_i^{(K)} = [x_i, x_{i-\tau}, \dots, x_{i-(K-1)\tau}]$ i $Y_i^{(L)} = [y_i, y_{i-\tau}, \dots, y_{i-(L-1)\tau}]$ są wektorami osadzonymi X i Y o wymiarach osadzenia odpowiednio K i L oraz okresie próbkowania τ . Licznik w wyrażeniu logarytmicznym w Równaniu 11 reprezentuje prawdopodobieństwo, że Y przyjmie określoną wartość (y_{i+h}) w przyszłości o h próbek, gdy znane są wcześniejsze wartości zarówno Y , jak i X . Mianownik w wyrażeniu logarytmicznym w Równaniu 11 reprezentuje prawdopodobieństwo, że Y przyjmie określoną wartość (y_{i+h}) w przyszłości o h próbek, gdy znane są jedynie wcześniejsze wartości Y . Gdy Y jest niezależne od X , licznik i mianownik są równe, a logarytm redukuje się do $\log(1) = 0$, co daje $T_{y \rightarrow x} = 0$. Gdy przewidywanie Y jest ulepszone przez uwzględnienie X , licznik jest większy od mianownika i $T_{x \rightarrow y} > 0$. Funkcje rozkładu prawdopodobieństwa można obliczyć za pomocą estymatora jądrowego gęstości (Bauer et al., 2007). Transfer Entropy nie jest ograniczona założeniem liniowości, dlatego może dawać dokładniejsze wyniki dla połączeń nieliniowych niż przyczynowość Grangera lub korelacja krzyżowa. Wadą tej metody jest fakt, że obliczenia funkcji rozkładu prawdopodobieństwa za pomocą estymatora jądrowego gęstości są wymagające obliczeniowo (wymagają ustawienia wielu parametrów) (Lindner et al., 2017a).

Transfer Entropy między dwiema zmiennymi X i Y może zostać obliczona zgodnie z poniższym:

$$T_{x \rightarrow y} = \text{transferEntropy}(x, y, \tau_x, \tau_y, N_p, C_{\text{thumb}}) \quad (12)$$

Gdzie x i y to szeregi czasowe w postaci jednowymiarowych wektorów, τ_x i τ_y są to opóźnienia w szeregach kolejno x i y , N_p to liczba równie rozmieszczonych punktów wzdłuż każdego wymiaru,

w którym szacuje się prawdopodobieństwa i C_{thumb} jest współczynnikiem (Bauer et al., 2007). Rozkład Gaussa jest używany do funkcji rozkładu prawdopodobieństwa w oryginalnym sformułowaniu algorytmu. Wynikiem funkcji jest $T_{x \rightarrow y}$, który jest miarą przepływu informacji pomiędzy dwiema zmiennymi x i y .

Metoda Transfer Entropy w naturalny sposób uwzględnia kierunkowość i dynamikę, ponieważ jest asymetryczna i opiera się na prawdopodobieństwach przejścia (Vicente et al., 2011). Główną zaletą tej metody jest to, iż nie zakłada ona żadnego konkretnego modelu dla wzajemnej relacji między zmiennymi. W ten sposób wrażliwość Transfer Entropy na wszystkie korelacje porządkowe staje się zaletą dla analiz eksploracyjnych (Kayser et al., 2009) w porównaniu z podejściami korelacji krzyżowej opartymi na modelach, przyczynowości Grangera lub częściowej koherencji ukierunkowanej. Bezmodelowe podejście jest więc kluczowe w sytuacjach, gdzie konieczne jest wykrycie nieznanymi, nieliniowych zależności. W takich przypadkach można oczekiwać, że dostarczy bardziej precyzyjnych i niezawodnych wyników. TE jest jednak ograniczona założeniem stacjonarności, co oznacza, że właściwości dynamiczne systemu nie mogą zmieniać się w zestawie danych użytych do analizy (Duan, 2014).

Bibliografia

- Ancona, N., Marinazzo, D. and Stramaglia, S. (2004), 'Radial basis function approach to nonlinear granger causality of time series', *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* **70**.
- Baccalá, L. A. and Sameshima, K. (2001), 'Partial directed coherence: A new concept in neural structure determination', *Biological Cybernetics* **84**.
- Bauer, M., Cox, J. W., Caveness, M. H., Downs, J. J. and Thornhill, N. F. (2007), 'Finding the direction of disturbance propagation in a chemical process using transfer entropy', *IEEE Transactions on Control Systems Technology* **15**.
- Bauer, M. and Thornhill, N. F. (2008), 'A practical method for identifying the propagation path of plant-wide disturbances', *Journal of Process Control* **18**.
- Boccaletti, S., Kurths, J., Osipov, G., Valladares, D. L. and Zhou, C. S. (2002), 'The synchronization of chaotic systems'.
- Butte, A. J. and Kohane, I. S. (2000), 'Mutual information relevance networks: functional genomic clustering using pairwise entropy measurements.', *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing* .
- Chen, Y., Rangarajan, G., Feng, J. and Ding, M. (2004), 'Analyzing multiple nonlinear time series with extended granger causality', *Physics Letters, Section A: General, Atomic and Solid State Physics* **324**.
- Duan, P. (2014), Information Theory-based Approaches for Causality Analysis with Industrial Applications, PhD thesis, University of Alberta. doi: .
- Granger, C. W. (2004), 'Time series analysis, cointegration, and applications'.
- Imbens, G. W. (2004), Nonparametric estimation of average treatment effects under exogeneity: A review, in 'Review of Economics and Statistics', Vol. 86.
- Kayser, A. S., Sun, F. T. and D'Esposito, M. (2009), 'A comparison of granger causality and coherency in fmri-based analysis of the motor system', *Human Brain Mapping* **30**.

- Kushwaha, N. and Lee, E. D. (2023), 'Discovering the mesoscale for chains of conflict', *PNAS Nexus* **2**.
- Landman, R., Kortela, J., Sun, Q. and Jämsä-Jounela, S. L. (2014), 'Fault propagation analysis of oscillations in control loops using data-driven causality and plant connectivity', *Computers and Chemical Engineering* **71**.
- Lindner, B., Auret, L. and Bauer, M. (2017a), Investigating the impact of perturbations in chemical processes on data-based causality analysis. part 1: Defining desired performance of causality analysis techniques, in 'IFAC-PapersOnLine', Vol. 50, p. 3269–3274.
- Lindner, B., Auret, L. and Bauer, M. (2017b), Investigating the impact of perturbations in chemical processes on data-based causality analysis. part 2: Testing granger causality and transfer entropy, in 'IFAC-PapersOnLine', Vol. 50, pp. 3275–3280.
- Lizier, J. T., Heinzle, J., Horstmann, A., Haynes, J. D. and Prokopenko, M. (2011), 'Multivariate information-theoretic measures reveal directed information structure and task relevant changes in fmri connectivity', *Journal of Computational Neuroscience* **30**.
- Mani, S. and Cooper, G. F. (2000), 'Causal discovery from medical textual data.', *Proceedings / AMIA ... Annual Symposium. AMIA Symposium*.
- Schreiber, T. (2000), 'Measuring information transfer', *Physical Review Letters* **85**.
- Steeg, G. V. and Galstyan, A. (2012), Information transfer in social media, in 'WWW'12 - Proceedings of the 21st Annual Conference on World Wide Web'.
- Vicente, R., Wibral, M., Lindner, M. and Pipa, G. (2011), 'Transfer entropy-a model-free measure of effective connectivity for the neurosciences', *Journal of Computational Neuroscience* **30**.
- Wiener, N. (1956), 'The theory of prediction', *Beckenbach EF (ed) Modern mathematics for engineers, McGraw-Hill, New York* **8**, 3269–3274.
- Yang, F., Duan, P., Shah, S. L. and Chen, T. (2014), *Capturing connectivity and causality in complex industrial processes*.
- Yu, W. and Yang, F. (2015), 'Detection of causality between process variables based on industrial alarm data using transfer entropy', *Entropy* **17**.