



PHÂN TÍCH DỮ LIỆU ĐIỆN NÃO ĐỒ BẰNG KỸ THUẬT XỬ LÝ TÍN HIỆU THỜI GIAN-TẦN SỐ

Cao Phương Thảo, Nguyễn Thị Hậu, Nguyễn Thanh Toàn
Trường Đại học Giao thông vận tải

Ngày nhận: 22/06/2016
Ngày sửa chữa: 16/08/2016
Ngày xét duyệt: 09/09/2016

Tóm tắt:

Nghiên cứu sự phụ thuộc giữa các chuỗi dữ liệu điện não đồ là công việc rất quan trọng để hiểu về hoạt động suy nghĩ của con người. Các nghiên cứu về sự phụ thuộc giữa các chuỗi dữ liệu điện não đồ sử dụng mô hình cấu trúc và tương quan đều giả sử cấu trúc phụ thuộc giữa các chuỗi dữ liệu không thay đổi theo thời gian. Trên thực tế sự phụ thuộc này thay đổi theo các phản ứng của con người. Bài báo giới thiệu kỹ thuật sử dụng kết hợp miền thời gian-tần số để phân tích dữ liệu điện não đồ dựa trên mô hình tự hồi quy. Kỹ thuật này cho phép ước lượng được sự thay đổi cấu trúc phụ thuộc theo thời gian dựa trên bộ lọc Kalman và liên kết có hướng cục bộ. Kết quả trên dữ liệu mô phỏng và thực tế cho thấy kỹ thuật này xác định chính xác các cấu trúc phụ thuộc thay đổi theo thời gian của dữ liệu điện não đồ.

Từ khóa: sóng điện não đồ, mô hình tự hồi quy, phụ thuộc có hướng.

1. Đặt vấn đề

Xác định sự phụ thuộc thay đổi theo thời gian giữa các cấu trúc nơ ron khác nhau là công việc quan trọng để hiểu về suy nghĩ của con người cũng như chẩn đoán sự bất thường trong các bệnh liên quan đến não [1]. Frison đưa ra khái niệm mô hình quan hệ động (DCM) [2] dựa trên các hệ thống trạng thái vào ra phi tuyến và xấp xỉ tương tác động để tính toán sự phụ thuộc giữa các chuỗi thời gian. Tuy nhiên kết quả của mô hình quan hệ động phải dựa trên cấu trúc kết nối có trước và sự phụ thuộc giữa các chuỗi dữ liệu không thay đổi theo thời gian. Khái niệm quan hệ nhân quả Granger [3] trong kinh tế lượng với ưu điểm là không cần xác định trước cấu trúc phụ thuộc, cũng đã được áp dụng để tính toán sự phụ thuộc giữa các chuỗi thời gian, đặc biệt là các tín hiệu sinh học. Các nhóm tác giả Roebroek, Goebel và cộng sự [4,5] cũng đã đề xuất phương pháp sử dụng mô hình tự hồi quy trong phân tích dữ liệu ảnh chụp não và sóng điện não đồ và đưa ra kết quả phụ thuộc giữa các miền trong não. Tuy nhiên các phương pháp này chưa đủ để kết luận sự phụ thuộc lẫn nhau giữa các tín hiệu vì các mô hình này dựa trên sự dự đoán giá trị của tín hiệu này dựa trên tín hiệu khác.

Trong bài báo này chúng tôi đề xuất sử dụng phương pháp liên kết có hướng cục bộ (partial directed coherence) [6], một trong những phương pháp miền thời gian-tần số dựa trên mô hình tự hồi quy đa biến theo khái niệm nhân quả Granger. Phương pháp này không yêu cầu phải giả sử có cấu trúc kết nối cho trước. Đặc biệt, phương pháp này có nhiều thuật toán để tính sự phụ thuộc thay đổi theo thời gian của các chuỗi dữ liệu. Mục tiêu của

bài báo là xây dựng được cấu trúc phụ thuộc thay đổi theo thời gian giữa các chuỗi dữ liệu điện não đồ. Kết quả trên dữ liệu mô phỏng và trên dữ liệu thực tế đã chứng minh phương pháp này khả thi trong phân tích dữ liệu điện não đồ.

2. Phương pháp

2.1. Mô hình tự hồi quy đa biến

Cho chuỗi thời gian y_n , mô hình tự hồi quy đa biến bậc p được định nghĩa như sau:

$$\begin{bmatrix} y_1[n] \\ \vdots \\ y_m[n] \end{bmatrix} = \sum_{k=1}^p A_k \begin{bmatrix} y_1[n-k] \\ \vdots \\ y_m[n-k] \end{bmatrix} + \begin{bmatrix} e_1[n] \\ \vdots \\ e_m[n] \end{bmatrix} \quad (1)$$

Với $[e_1, \dots, e_m]^T = e$ là vector nhiễu với ma trận hiệp phương sai đường chéo $\Sigma_e = \langle ww^T \rangle = \text{diag}\{\lambda_{rr}^2\}$.

Ma trận A_k được định nghĩa như sau:

$$A_k = \begin{bmatrix} a_{11}^k & \dots & a_{1m}^k \\ \vdots & \ddots & \vdots \\ a_{m1}^k & \dots & a_{mm}^k \end{bmatrix} \quad (2)$$

Với $k = 1, \dots, p$. Tham số a_{ij}^k phản ánh quan hệ tuyến tính giữa kênh i và kênh j với độ trễ k . Nếu chuỗi dữ liệu là cân bằng (stationary), bậc p của mô hình tự hồi quy có thể ước lượng bằng các phương pháp tiêu chuẩn Akaike Information Criterion (AIC) [7] hoặc Schwarz's Bayesian Criterion (SBC) [8].

2.2. Ước lượng các tham số

Hầu hết các phương pháp ước lượng tham số của mô hình tự hồi quy đa biến đều dựa trên giả sử chuỗi dữ liệu là ổn định. Trên thực tế, tất cả các dữ liệu sinh học đều không ổn định, cấu trúc phụ thuộc

thay đổi theo thời gian. Bài báo này giới thiệu hai phương pháp ước lượng tham số của mô hình tự hồi quy đa biến với dữ liệu không ổn định là phương pháp cửa sổ thời gian ngắn [9] và phương pháp lọc Kalman [10].

Phương pháp cửa sổ thời gian ngắn (short-window)

Ở phương pháp này, các chuỗi dữ liệu được chia thành các cửa sổ thời gian đủ ngắn chom nhau và coi chuỗi dữ liệu trong các cửa sổ đó là ổn định. Với mỗi chuỗi dữ liệu trong từng cửa sổ, áp dụng mô hình tự hồi quy tuyến tính cho các chuỗi dữ liệu đó và ước lượng các tham số của mô hình tự hồi quy. Kết quả là một mạng các tham số của từng chuỗi dữ liệu ổn định. Kết hợp các tham số theo thứ tự thời gian, ta có thể coi đây là các tham số thay đổi của mô hình tự hồi quy trên chuỗi dữ liệu không ổn định. Tuy nhiên với phương pháp này việc chọn kích thước cửa sổ sao cho phù hợp là điều không dễ dàng.

Phương pháp lọc Kalman

Từ mô hình tự hồi quy không ổn định, công thức (1) chuyển thành

$$\begin{aligned} x_t &= \text{Vec}(a'_{k,t}) \\ C_t &= (I_d \otimes y'_{t-1}, \dots, y'_{t-p-1}) \end{aligned} \quad (3)$$

Với Vec là toán tử biểu diễn các hệ số $a_{k,t}$ dưới dạng vector, I_d là ma trận đơn vị và \otimes là phép toán Kronecker.

Mô hình tự hồi quy có thể biểu diễn dưới dạng không gian trạng thái như sau:

$$\begin{aligned} y_t &= C_t x_t + v_t \\ x_{t+1} &= A x_t + w_t \end{aligned} \quad (4)$$

Với x_t là các hệ số của mô hình tự hồi quy, w_t là nhiễu với trung bình là không và phương sai Q . A là ma trận chuyển trạng thái và C_t là ma trận quan sát. Giả sử trạng thái khởi tạo theo phân phối Gaussian với trung bình là μ và phương sai là Σ . Ở mô hình này, các tham số cần ước lượng là $\Theta = \{A, Q, R, \mu, \Sigma\}$.

Với chuỗi dữ liệu $Y_{1:T}$, áp dụng bộ lọc Kalman để tính các hệ số của mô hình tự hồi quy, ta cần tính kỳ vọng của trạng thái và phương sai của sai số

$$x_{t|k} = E(x_t | Y_{1:T}) \quad (5)$$

$$P_{t|t} = E\{(x_t - x_{t|k})(x_t - x_{t|k})' | Y_{1:T}\} \quad (6)$$

Để tính trạng thái $x_{t|t}$ tại thời điểm t , ta cần tính trạng thái $x_{t|t-1}$ dựa vào trạng thái trước đó $x_{t-1|t-1}$. Cho trước chuỗi dữ liệu y , ta áp dụng bộ lọc Kalman tiến để duyệt các hệ số cho đến đến hệ số x_T , sau đó áp dụng làm mịn Kalman ngược để tính lại x_t .

Để tính các tham số, ta sử dụng giải thuật cực đại hóa kỳ vọng, giải thuật này thực hiện cực đại hóa kỳ vọng của chuỗi dữ liệu Y với các hệ số ẩn $x_{t|T}$. Bước kỳ vọng tính log của khả năng:

$$F = E_{X|Y}[\log P(Y_{1:T}, X_{1:T} | \Theta)] \quad (7)$$

Xét Y là dữ liệu quan sát còn trạng thái X là các biến ẩn. Mỗi tham số trong bước cực đại hóa được tính bằng cách tính kỳ vọng của log khả năng:

$$\hat{R} = \frac{1}{T} \sum_{t=1}^T (y_t y'_t - 2C_t \hat{x}_t y'_t + C_t P_t C'_t) \quad (8)$$

$$\hat{A} = \left(\sum_{t=2}^T P_{t,t-1} \right) \left(\sum_{t=2}^T P_{t-1} \right)^{-1} \quad (9)$$

$$\hat{Q} = \frac{1}{T-1} \left(\sum_{t=2}^T P_t - \hat{A} \sum_{t=2}^T P_{t-1,t} \right) \quad (10)$$

$$\hat{\mu} = \hat{x}_t \quad (11)$$

$$\hat{\Sigma} = P_t - x_t x'_t \quad (12)$$

Sau khi tính được các tham số của không gian trạng thái, ta thay ngược lại công thức (1) để tính được các hệ số của mô hình tự hồi quy bằng cách đảo ngược lại vector trạng thái x .

2.2. Ước lượng phụ thuộc giữa các chuỗi dữ liệu

Liên kết có hướng cục bộ (PDC) là phương pháp phân tích sự phụ thuộc có hướng giữa các tín hiệu theo ngữ cảnh phụ thuộc Granger. Để quan sát sự phụ thuộc biến đổi theo thời gian giữa các chuỗi dữ liệu điện não đồ, chúng tôi mở rộng phương pháp liên kết có hướng cục bộ theo thời gian (tvPDC) dựa trên các tham số biến đổi lấy từ các tham số của mô hình tự hồi quy. Để tính tvPDC, ta cần chuyển các tham số của mô hình tự hồi quy sang miền tần số bằng biến đổi Fourier.

$$A(f, t) = I - \sum_{k=1}^p a(t-k) \exp - ifk \quad (13)$$

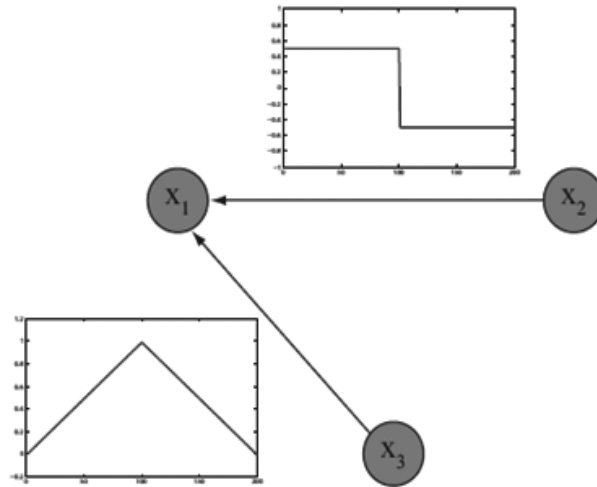
Liên kết từng phần có hướng từ x_i sang x_j được tính như sau:

$$PDC_{i \rightarrow j}(f, t) = \frac{|A_{ij}(f, t)|}{\sqrt{\sum_{l=1}^K |A_{lj}(f, t)|^2}} \quad (14)$$

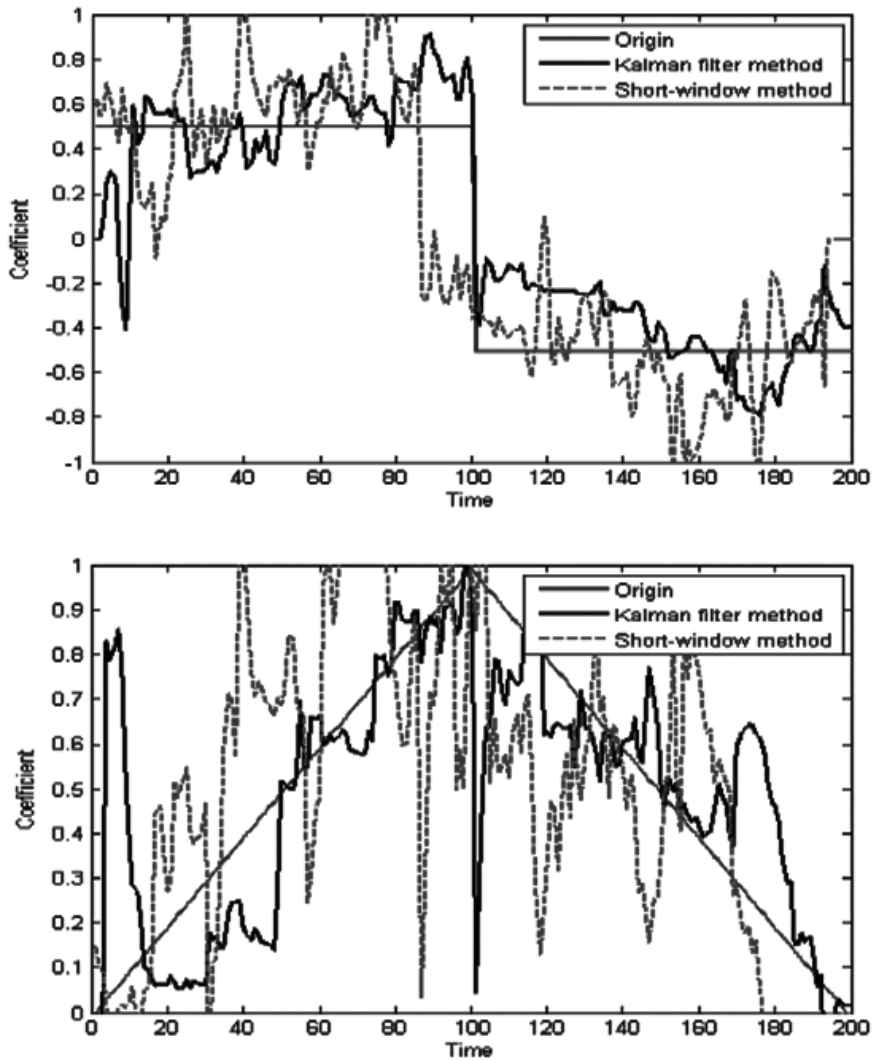
3. Kết quả

3.1. Dữ liệu mô phỏng

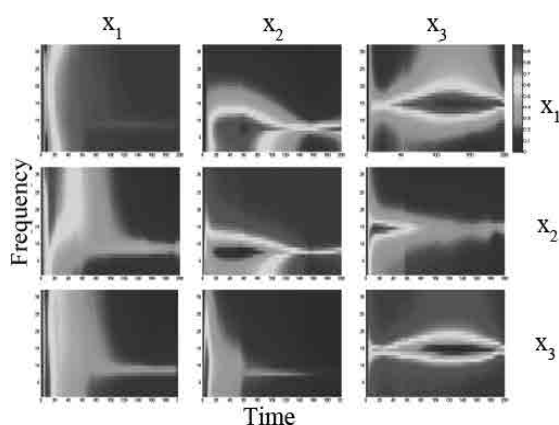
Để minh họa tính đúng đắn của hai phương pháp, bài báo sử dụng dữ liệu mô phỏng mô hình tự hồi quy ba biến bậc 2 được sinh ra từ công thức (1) với tham số thay đổi như trong Hình 1. Trong mô hình này, chuỗi dữ liệu x_2 ảnh hưởng đến x_1 với giá trị tăng ở giai đoạn đầu và giảm đột ngột ở giai đoạn sau. Chuỗi dữ liệu x_3 ảnh hưởng đến chuỗi dữ liệu x_1 với giá trị tăng dần giai đoạn đầu và giảm dần ở giai đoạn sau. Các tham số của mô hình tự hồi quy ước lượng theo hai phương pháp cửa sổ thời gian và bộ lọc Kalman được minh họa trong Hình 2. Từ kết quả ta thấy các tham số ước lượng bằng bộ lọc Kalman chính xác hơn phương pháp cửa sổ thời gian, đặc biệt là ở những vị trí tham số thay đổi. Kết quả của phương pháp cửa sổ thời gian kém hơn là do cửa sổ thời gian phải lớn hơn bậc của mô hình tự hồi quy.



Hình 1. Sự phụ thuộc giữa các chuỗi dữ liệu với tham số thay đổi



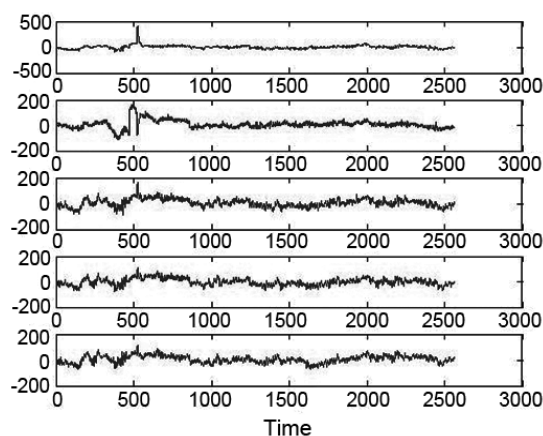
Hình 2. Kết quả ước lượng tham số theo hai phương pháp



Hình 3. Kết quả sự phụ thuộc giữa các chuỗi dữ liệu được tính bằng liên kết có hướng cục bộ. Phần tử (1,2) chỉ ra sự ảnh hưởng của chuỗi dữ liệu x_2 lên x_1 . Phần tử (1,3) chỉ ra sự ảnh hưởng của chuỗi dữ liệu x_2 lên x_1

3.2. Dữ liệu điện não đồ

Dữ liệu điện não đồ minh họa trong bài báo này được lấy từ phần mềm EEGLAB. Dữ liệu thu từ não người với hai sự kiện, một sự kiện khi người tham gia thí nghiệm nhìn thấy hình vuông màu xanh ở các vị trí khác nhau trên màn hình và một sự kiện người đó tác động lại. Trong dữ liệu này, người tham gia thí nghiệm sẽ quan sát hình vuông xem xuất hiện ở đâu để bấm nút tương ứng với vị trí hình vuông xuất hiện trên màn hình. Thí nghiệm này nhằm minh họa cơ chế nhìn và ra lệnh của não. Hình 4 chỉ ra các tín hiệu điện não đồ thu được ở 5 kênh thị giác và điều khiển.

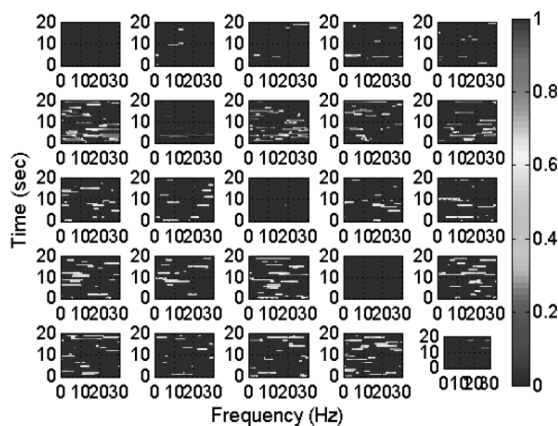


Hình 4. Dữ liệu điện não đồ 5 kênh

Tài liệu tham khảo

- [1]. Rieder, M. K., Rahm, B., Williams, J. D., and Kaiser, J., (2011), *Human-band Activity and Behavior*, International Journal of Psychophysiology, 79:39-48, 2011.
- [2]. K.J. Friston, L. Harisson, and W.D. Penny, "Dynamic Causal Modelling," Neuroimage, Vol. 19, pp. 1273-1302, 2003

Chuỗi dữ liệu điện não đồ 5 kênh được đưa vào mô hình tự hồi quy 5 biến và được chuyển sang không gian trạng thái như ở công thức (4). Các tham số được tính toán bằng phương pháp lọc Kalman theo công thức (8-12) và chuyển lại mô hình tự hồi quy đa biến theo thời gian. Các tham số này được đưa đến đầu vào của liên kết có hướng cục bộ (13-14). Sự phụ thuộc giữa các chuỗi dữ liệu điện não đồ được chỉ ra trong hình (5).



Hình 5. Sự phụ thuộc theo thời gian của 5 kênh dữ liệu điện não đồ

Từ hình trên ta thấy phần tử (2,1), (2,3), (2,4), (2,5), (3,1), (3,2), (3,4), (3,5), (4,1), (4,2), (4,3), (4,5), (5,1), (5,2), (5,3), (5,4), (5,5) chỉ ra sự phụ thuộc giữa các cặp kênh đó theo thời gian ở những tần số khác nhau. Dựa vào sự phụ thuộc này ta có thể xây dựng các cấu trúc phụ thuộc giữa các vùng trong não tương ứng với những điều kiện nhận thức khác nhau.

4. Kết luận

Bài báo trình bày phương pháp để tính toán sự phụ thuộc theo thời gian giữa các chuỗi dữ liệu điện não đồ. Phương pháp này giải quyết được hạn chế của việc phải giả sử các chuỗi dữ liệu là ổn định, dẫn đến cấu trúc phụ thuộc cố định theo thời gian. Ngoài ra phương pháp này còn có ưu điểm hơn phương pháp cửa sổ thời gian ngắn khi có những thay đổi lớn về cấu trúc phụ thuộc.

- [3]. Granger, C. W. J., “*Investigating Causal Relations by Econometric Models and Crossspectral Methods*”. *Econometrica*, 37(3):424-438, 1969
- [4]. Roebroeck, A., Formisano, E., and Goebel, R., “*Mapping Directed Influence over the Brain using Granger Causality and Fmri*”. *NeuroImage*, 25(1):230-242, 2005.
- [5]. Goebel R, Roebroeck A, Kim DS, Formisano E, “*Investigating Directed Cortical Interactions in Time-resolved fMRI Data using Vector Autoregressive Modeling and Granger Causality Mapping*”. *Magn Reson Imaging* 21:1251–1261, 2003.
- [6]. L. Baccala, and K. Sameshima, “*Partial Directed Coherence: A New Concept in Neural Structure Determination*,” *Biol. Cybern.*, Vol. 84, pp. 463-474, 2001.
- [7]. Akaike, H., *Information Theory and an Extension of the Maximum Likelihood Principle*, 2nd International Symposium on Information Theory, B. N. Petrov and F. Csaki (eds.), Akademiai Kiado, Budapest, 267 – 281, 1973.
- [8]. Schwarz, G., “*Estimating the Dimension of a Model*”, *Annals of Statistics*, 6, 461–464, 1978.
- [9]. Mingzhou Ding, Steven L. Bressler, Weiming Yang, Hualou Liang, “*Short-window Spectral Analysis of Cortical Event-related Potentials by Adaptive Multivariate Autoregressive Modeling: Data Preprocessing, Model Validation, and Variability Assessment*”, *Biological Cybernetics*, Springer-Verlag, vol. 83, pp 35-45, 2000.
- [10]. S.S. Haykin, “*Kalman Filtering and Neural Networks*,” Wiley Chichester, 2001.

EEG ANALYSIS USING TIME-FREQUENCY SIGNAL PROCESSING TECHNIQUE

Abstract:

Analyzing the interaction among electroencephalogram (EEG) data is an important step to understand the brain process. There have been many studies of directional effective influence between regions of the brain using EEG. However, most of the previous studies are based on the assumption of time-invariant connectivity structure, which is insufficient to understand the change of the connectivity during cognitive tasks. In this paper, we introduce the time-frequency method to analyzing the eeg data using autoregressive model. The method allows the examination of the dynamic interactions among brain regions during cognitive tasks using Kalman filter and Partial Directed Coherence. The result of the simulation study and the real data indicates that this approach is effective to estimate the time-varying connectivity among EEG data.

Keywords: eeg, autoregressive model, directed influence.