



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0029952

(51)⁷ G06K 9/00

(13) B

(21) 1-2018-03321

(22) 27/07/2018

(45) 25/11/2021 404

(43) 25/10/2018 367A

(73) Tập đoàn Công nghiệp - Viễn thông Quân Đội (VN)

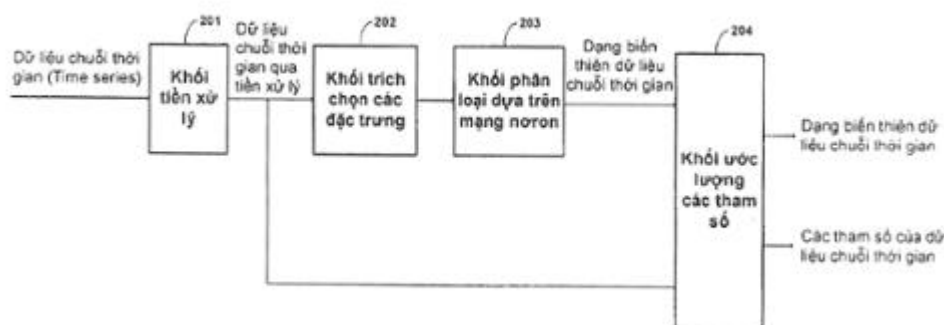
Số 1 đường Trần Hữu Dực, phường Mỹ Đình 2, quận Nam Từ Liêm, thành phố Hà Nội

(72) Nguyễn Phan Khánh Hà (VN); Nguyễn Quý Hà (VN); Nguyễn Thái Bình (VN); Khương Duy (VN).

(74) Công ty TNHH Tư vấn Quốc Dân (NACILAW)

(54) HỆ THỐNG VÀ PHƯƠNG PHÁP TỰ ĐỘNG PHÂN LOẠI, ƯỚC LƯỢNG CÁC DẠNG BIẾN THIÊN CỦA DỮ LIỆU CHUỖI THỜI GIAN

(57) Sáng chế đề xuất một hệ thống và phương pháp tự động phân loại, ước lượng các dạng biến thiên của dữ liệu chuỗi thời gian sử dụng mạng nơron dựa trên các đặc trưng được trích chọn từ chuỗi dữ liệu đầu vào. Cụ thể là, hệ thống và phương pháp tự động phân loại, ước lượng tham số dạng biến thiên chuỗi thời gian sử dụng các khối: khối tiền xử lý, khối trích chọn các đặc trưng, khối phân loại dựa trên mạng nơron truyền thẳng và khối ước lượng các tham số; được thể hiện qua các bước: bước 1: thực hiện lọc trung vị đối với dữ liệu chuỗi thời gian đầu vào; bước 2: trích chọn các đặc trưng của các dạng biến thiên dữ liệu chuỗi thời gian; bước 3: xác định dạng biến thiên dữ liệu chuỗi thời gian; bước 4: ước lượng các đặc trưng ứng với dạng biến thiên dữ liệu chuỗi thời gian.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập tới hệ thống và phương pháp tự động phân loại, ước lượng các dạng biến thiên của dữ liệu chuỗi thời gian (time series). Cụ thể, hệ thống và phương pháp tự động phân loại, ước lượng các dạng biến thiên của dữ liệu chuỗi thời gian được đề cập trong sáng chế này sử dụng mạng nơron (neural network) dựa trên trích chọn các đặc trưng từ dữ liệu chuỗi thời gian đầu vào.

Tình trạng kỹ thuật của sáng chế

Dữ liệu chuỗi thời gian là một chuỗi các điểm dữ liệu rời rạc được đo theo những khoảng thời gian liên tiếp nhau với tần suất thời gian thống nhất. Khái niệm này được sử dụng trong nhiều lĩnh vực khác nhau như xử lý tín hiệu, khai phá dữ liệu, tài chính, kinh tế lượng, khí tượng... Một số ví dụ của dữ liệu chuỗi thời gian như: doanh số bán hàng của công ty trong một giai đoạn, biến động giá cổ phiếu của một công ty, độ cao của thủy triều trong ngày; nhiệt độ, lượng mưa ghi nhận tại một trạm quan trắc khí tượng; công suất tiêu thụ điện năng trong một nhà máy, chu kỳ lặp các xung tín hiệu liên tiếp nhau của một đài radar...

Dữ liệu chuỗi thời gian thường có xu hướng biến thiên tuân theo một dạng nhất định. Hình 1 a đến hình 1g minh họa các dạng biến thiên của dữ liệu chuỗi thời gian thường gặp bao gồm:

- Dạng cố định (Constant): dữ liệu chuỗi thời gian là hằng số.
- Dạng biến thiên tuyến tính tăng (Sliding Up): dữ liệu chuỗi thời gian biến đổi tuyến tính tăng theo một đường dốc xác định.
- Dạng biến thiên tuyến tính giảm (Sliding Down): dữ liệu chuỗi thời gian biến đổi tuyến tính giảm theo một đường dốc xác định.
- Dạng biến thiên ngẫu nhiên (Jitter): dữ liệu chuỗi thời gian biến đổi ngẫu nhiên quanh một giá trị trung bình trong một độ lệch xác định.
- Dạng biến thiên so le (Staggered): dữ liệu chuỗi thời gian biến đổi tuần tự theo qua một số các mức xác định khác nhau và lặp lại.

- Dạng biến thiên bậc thang (Dwell&Switch): dữ liệu chuỗi thời gian biến đổi qua từng bậc (trong mỗi bậc, giá trị dữ liệu chuỗi thời gian là cố định).
- Dạng biến thiên theo chu kỳ (Periodic): dữ liệu chuỗi thời gian biến đổi theo một dạng tuần hoàn (sin, cos hoặc hàm mũ, logarit...).

Việc tự động xác định dạng biến thiên của dữ liệu chuỗi thời gian và ước lượng các tham số đặc trưng của dữ liệu chuỗi thời gian có ý nghĩa không chỉ trong việc phân tích, thống kê mà còn là cơ sở cho việc đưa ra các dự báo trong tương lai. Các phương pháp tự động phân loại các dạng biến thiên dữ liệu chuỗi thời gian truyền thống thường theo hai hướng:

- Dựa trên cây quyết định (Decision-tree);
- Dựa trên học máy với dữ liệu đầu vào trực tiếp là dữ liệu chuỗi thời gian.

Phương pháp phân loại dựa trên cây quyết định có hạn chế cơ bản là đòi hỏi việc thiết lập các ngưỡng quyết định. Trong nhiều trường hợp, việc thiết lập các ngưỡng đòi hỏi nhân sự có kinh nghiệm về dữ liệu và tốn thời gian do số lượng ngưỡng cần thiết lập lớn và quá trình tối ưu giá trị ngưỡng là phức tạp (đòi hỏi thử nghiệm và tinh chỉnh). Đồng thời hiệu năng của phương pháp này bị ảnh hưởng bởi mức độ nhiễu của dữ liệu, khi môi trường thay đổi sẽ mất thời gian và công sức để điều chỉnh và thiết lập lại các ngưỡng phù hợp.

Phương pháp phân loại dựa trên học máy với dữ liệu đầu vào trực tiếp là dữ liệu chuỗi thời gian mặc dù khắc phục được yêu cầu thiết lập ngưỡng của phương pháp cây quyết định nhưng lại đòi hỏi độ dài chuỗi đầu vào là cố định. Tuy nhiên, trong thực tế, độ dài dữ liệu chuỗi thời gian có thể thay đổi, dẫn đến tính linh hoạt của phương pháp này là hạn chế. Bên cạnh đó, hiệu năng của phương pháp này cũng bị ảnh hưởng bởi mức độ nhiễu của dữ liệu vào.

Trong sáng chế này, tác giả đề xuất một hệ thống và phương pháp để tự động phân loại và ước lượng các dạng biến thiên dữ liệu chuỗi thời gian sử dụng mạng nơron dựa trên các đặc trưng được trích chọn từ dữ liệu chuỗi thời gian đầu vào.

Bản chất kỹ thuật của sáng chế

Mục đích thứ nhất của sáng chế là đề xuất hệ thống tự động phân loại và ước lượng các dạng biến thiên của dữ liệu chuỗi thời gian, hệ thống này bao gồm các khối: khối tiền xử lý, khối trích chọn các đặc trưng, khối phân loại dựa trên mạng nơron, và khối ước lượng các tham số.

Mục đích thứ hai của sáng chế là đề xuất phương pháp tự động phân loại và ước lượng các dạng biến thiên của dữ liệu chuỗi thời gian, phương pháp này bao gồm các bước như sau: bước 1: thực hiện lọc trung vị đối với dữ liệu chuỗi thời gian đầu vào; bước 2: trích chọn các đặc trưng của các dạng biến thiên dữ liệu chuỗi thời gian; bước 3: xác định dạng biến thiên dữ liệu chuỗi thời gian; bước 4: ước lượng các đặc trưng ứng với dạng biến thiên dữ liệu chuỗi thời gian. Kết quả đầu ra của hệ thống là dạng biến thiên dữ liệu chuỗi thời gian và các tham số tương ứng đã ước lượng.

Trong sáng chế này, do dựa trên việc sử dụng mạng nơron đã được huấn luyện với các đặc trưng được trích chọn từ dữ liệu chuỗi thời gian để tự động phân loại dạng biến thiên nên giải pháp đề xuất đơn giản hóa quá trình thực thi và không đòi hỏi việc thiết lập ngưỡng cho việc phân loại như đối với các phương pháp dựa trên cây quyết định; giảm thiểu thời gian và nhân lực có kinh nghiệm thực hiện việc thiết lập rất nhiều ngưỡng phân loại khác nhau.

Trong sáng chế này, giải pháp đề xuất thực hiện bước tiền xử lý (lọc trung vị) với dữ liệu đầu vào để hạn chế ảnh hưởng của nhiễu lên dữ liệu vào trước khi thực hiện các bước xử lý tiếp theo nên có khả năng thích ứng tốt, nâng cao độ chính xác phân loại trong môi trường có nhiễu cao.

Trong sáng chế này, do dựa trên việc sử dụng mạng nơron với đầu vào là các đặc trưng được trích chọn nên giải pháp đề xuất có thể hoạt động được với dữ liệu chuỗi thời gian đầu vào có độ dài chuỗi thay đổi do việc trích chọn các đặc trưng sử dụng khối trích chọn các đặc trưng được đề xuất hoàn toàn độc lập, không phụ thuộc vào độ dài chuỗi dữ liệu thời gian đầu vào. Đây là đặc tính làm cho giải pháp đề xuất có tính linh hoạt hơn so với các phương pháp phân loại dựa trên học máy với dữ liệu chuỗi thời gian đầu vào thô do phương pháp này yêu cầu dữ liệu chuỗi thời gian đầu vào cần có độ dài cố định.

Trong sáng chế này, giải pháp đề xuất có tính tổng quát cao, có thể áp dụng cho bất kỳ loại dữ liệu chuỗi thời gian trong các lĩnh vực khác nhau như kinh tế, tài chính, tín hiệu thông tin, tín hiệu radar, thống kê... có đặc tính biến thiên theo một dạng xác định.

Mô tả vắn tắt các hình vẽ

Hình 1a: hình ảnh minh họa dạng cố định của dữ liệu chuỗi thời gian.

Hình 1b: hình ảnh minh họa dạng biến thiên tuyến tính tăng của dữ liệu chuỗi thời gian.

Hình 1c: hình ảnh minh họa dạng biến thiên tuyến tính giảm của dữ liệu chuỗi thời gian.

Hình 1d: hình ảnh minh họa dạng biến thiên ngẫu nhiên của dữ liệu chuỗi thời gian.

Hình 1e: hình ảnh minh họa dạng biến thiên sole của dữ liệu chuỗi thời gian.

Hình 1f: hình ảnh minh họa dạng biến thiên bậc thang của dữ liệu chuỗi thời gian.

Hình 1g: hình ảnh minh họa dạng biến thiên theo chu kỳ của dữ liệu chuỗi thời gian.

Hình 2: hình ảnh minh họa hệ thống tự động phân loại và ước lượng các dạng biến thiên của dữ liệu chuỗi thời gian.

Hình 3: sơ đồ minh họa phương pháp tự động phân loại và ước lượng các dạng biến thiên của dữ liệu chuỗi thời gian.

Hình 4a: hình ảnh minh họa dữ liệu chuỗi thời gian đầu vào (với nhiễu do các điểm dữ liệu sai lệch).

Hình 4b: hình ảnh minh họa dữ liệu chuỗi thời gian sau khi qua khối tiền xử (lọc trung vị).

Hình 5: hình ảnh minh họa cấu trúc khối trích chọn các đặc trưng.

Hình 6: hình ảnh minh họa cấu trúc khối trích chọn các đặc trưng dựa trên hàm tự tương quan.

Hình 7: hình ảnh minh họa cấu trúc khối trích chọn các đặc trưng dựa trên các chuỗi mã.

Hình 8: hình ảnh minh họa kiến trúc mạng nơron sử dụng cho phân loại dạng biến thiên dữ liệu chuỗi thời gian.

Hình 9: hình ảnh minh họa cấu trúc khối ước lượng các tham số.

Hình 10: hình ảnh minh họa cấu trúc khối ước lượng tham số cho dạng biến thiên nhiều mức.

Hình 11: hình ảnh minh họa quá trình xử lý ở khối ước lượng các tham số.

Hình 12a: hình ảnh minh họa một dạng biến thiên dữ liệu chuỗi thời gian sau tiền xử lý.

Hình 12b: sơ đồ minh họa dạng đồ thị tần suất (histogram) tương ứng với một dạng biến thiên dữ liệu chuỗi thời gian sau tiền xử lý.

Mô tả chi tiết sáng chế

Để đạt được mục đích thứ nhất của sáng chế là đề xuất một hệ thống tự động phân loại và ước lượng các dạng biến thiên của dữ liệu chuỗi thời gian, hệ thống này bao gồm các khối: khối tiền xử lý, khối trích chọn các đặc trưng, khối phân loại dựa trên mạng nơron truyền thẳng và khối ước lượng các tham số. Cụ thể:

- Khối tiền xử lý

Dữ liệu chuỗi thời gian $\{p[n]\}_{n=0}^{N-1}$ đầu vào của hệ thống là một chuỗi các điểm dữ liệu rời rạc, gồm N mẫu được đo theo những khoảng thời gian liên tiếp cách đều nhau (lưu ý là giữa các lần phân tích dạng biến thiên dữ liệu chuỗi thời gian, độ dài của chuỗi dữ liệu thời gian này không bắt buộc phải đồng nhất, nghĩa là N có thể thay đổi được).

Tham chiếu hình 2, khối tiền xử lý 201 có chức năng giảm ảnh hưởng của nhiễu đối với dữ liệu chuỗi thời gian đầu vào $\{p[n]\}$. Khối tiền xử lý 201 là một bộ

lọc trung vị (median filter) có kích thước cửa sổ là K mẫu. Dữ liệu đầu vào $\{p[n]\}$ được lọc sử dụng bộ lọc trung vị. Đầu ra của khối tiền xử lý là dữ liệu chuỗi thời gian $\{p'[n]\}$ đã được lọc bỏ nhiễu. Tham chiếu hình 4a và hình 4b, hình ảnh minh họa tác dụng của bộ lọc trung vị trong giảm ảnh hưởng của nhiễu đối với dữ liệu chuỗi thời gian đầu vào (ví dụ cho dạng biến thiên tuyến tính giảm).

- Khối trích chọn các đặc trưng

Tham chiếu hình 2, khối trích chọn các đặc trưng 202 có chức năng phân tích và trích xuất 6 đặc trưng cho dạng biến thiên dữ liệu chuỗi thời gian từ dữ liệu chuỗi thời gian đầu vào sau tiền xử lý. Đầu vào của khối trích chọn các đặc trưng 202 là dữ liệu chuỗi thời gian đã qua khối tiền xử lý 201. Đầu ra của khối này là bộ 6 đặc trưng được trích chọn của dữ liệu chuỗi thời gian. Tham chiếu hình 5, khối trích chọn đặc trưng gồm có ba khối con:

- Khối trích chọn các đặc trưng dựa trên hàm tự tương quan 501: khối này trích chọn ra 2 đặc trưng là tỷ lệ giữa độ biến thiên của dữ liệu chuỗi thời gian với giá trị trung bình và tỷ lệ giữa giá trị hàm tự tương quan với trễ 0 mẫu và trễ 1 mẫu.
- Khối trích chọn đặc trưng ước lượng số nhóm trong dữ liệu chuỗi thời gian 502: khối này trích chọn ra một đặc trưng là số lượng nhóm khác nhau trong dữ liệu chuỗi thời gian.
- Khối trích chọn các đặc trưng dựa trên chuỗi mã ký tự hóa 503: khối này trích chọn ra 3 đặc trưng là chỉ số đo mức độ hỗn loạn entropy (entropy) chuẩn hóa của đồ thị tần suất của chuỗi mã hóa dữ liệu chuỗi thời gian, độ nhọn (kurtosis) của phân bố của chuỗi mã của dữ liệu chuỗi thời gian và mômen bậc hai chuẩn hóa (normalized second-order moment) của chuỗi mã của chuỗi sai phân bậc một dữ liệu chuỗi thời gian (chuỗi sai phân bậc một của dữ liệu chuỗi thời gian $\{d[n]\}$ được định nghĩa $d[n] = p'[n+1] - p'[n]; n = 0, \dots, N-2$).

Cấu tạo cụ thể của ba khối con này như sau:

- a. Khối trích chọn các đặc trưng dựa trên hàm tự tương quan

Cấu trúc của khối này được minh họa trong hình 6 gồm có hai khối con là: khối tính toán hàm tự tương quan 601 và khối tính toán 2 đặc trưng dựa trên hàm tự tương quan 602.

Tham chiếu hình 6, đầu vào của khối là dữ liệu chuỗi thời gian đã qua khối tiền xử lý 201. Khối tính toán hàm tự tương quan 601 thực hiện tính toán hàm tự tương quan của chuỗi theo công thức (CT1):

$$A_p[k] = \frac{1}{N} \sum_{n=0}^{N-k-1} (p'[n] - \mu_p)(p'[n+k] - \mu_p) \quad (CT1)$$

trong đó, μ_p là giá trị trung bình của chuỗi $\{p'[n]\}$.

Bản chất của việc sử dụng hàm tự tương quan là để đánh giá mức độ giống nhau giữa chuỗi thời gian và các phiên bản dịch thời gian (time-shift version) của nó. Dựa trên hàm tự tương quan, khối tính toán 2 đặc trưng dựa trên hàm tự tương quan 602 tính toán hai đặc trưng là tỷ lệ giữa độ biên thiên của chuỗi thời gian với giá trị trung bình và tỷ lệ giữa giá trị hàm tự tương quan với trễ 0 mẫu và trễ 1 mẫu theo công thức (CT2) và (CT3). Hai công thức này được tham khảo dựa trên nội dung trong bài báo “Phát hiện dạng PRI sử dụng nhóm mức và hàm tính tự tương quan” (“PRI modulation type recognition using level clustering and autocorrelation”) của M. Ahmadi và K. Mohamedpour công bố trên tạp chí “Xử lý tín hiệu Hoa Kỳ” (“American Journal of Signal Processing”) xuất bản năm 2012 (từ trang 83 đến trang 91):

$$f_1 = \frac{\sqrt{A_p[0]}}{\mu_p} \quad (CT2)$$

$$f_2 = 1 - \frac{A_p[1]}{A_p[0]} \quad (CT3)$$

Đầu ra của khối là hai đặc trưng được trích chọn cho dữ liệu chuỗi thời gian. Đặc trưng đầu tiên giúp phân biệt dạng dữ liệu chuỗi thời gian cố định (biến thiên rất nhỏ) với các dạng biến thiên khác, trong khi đặc trưng thứ hai giúp phân biệt dạng biến thiên dữ liệu chuỗi thời gian so le (khác biệt giữa các mẫu liên tiếp nhau thường lớn) với các dạng biến thiên khác.

b. Khối trích chọn đặc trưng ước lượng số nhóm trong dữ liệu chuỗi thời gian

Mục đích của khối là trích chọn đặc trưng về số lượng nhóm khác nhau trong dữ liệu chuỗi thời gian. Đầu vào của khối là dữ liệu chuỗi thời gian đã qua khối tiền xử lý. Khối trích chọn đặc trưng ước lượng số nhóm trong dữ liệu chuỗi thời gian 502 để ước lượng số nhóm khác nhau trong dữ liệu chuỗi thời gian theo công thức (CT4):

$$f_3 = \text{Số lượng nhóm khác nhau trong chuỗi } \{p'[n]\} \quad (\text{CT4})$$

Bản chất của đặc trưng này là giúp phân biệt các dạng biến thiên liên tục giữa các mẫu với các dạng biến thiên gồm nhiều mức giá trị rời rạc (dạng biến thiên so le và dạng biến thiên bậc thang).

c. Khối trích chọn các đặc trưng dựa trên chuỗi mã ký tự hóa

Mục đích của khối này là trích chọn 3 đặc trưng biểu diễn cho dữ liệu chuỗi thời gian thông qua việc sử dụng kỹ thuật ký tự hóa để tạo các chuỗi mã đối với dữ liệu chuỗi thời gian và chuỗi sai phân bậc một của dữ liệu chuỗi thời gian. Ý tưởng sử dụng kỹ thuật ký tự hóa để tạo các chuỗi mã được tham khảo dựa trên nội dung bài báo “Đánh giá về (kỹ thuật) phân tích ký tự cho dữ liệu thực nghiệm” (“A review of symbolic analysis of experimental data”) của C. S. Daw, C. E. A. Finney, E. R. Tracy công bố trên tạp chí “Đánh giá các công cụ khoa học” (Review of Scientific Instruments) xuất bản năm 2003 (từ trang 915 đến 930), theo đó, đối với một chuỗi dữ liệu thời gian việc sử dụng một kỹ thuật ký tự hóa để chuyển thành một chuỗi mã mang theo các đặc trưng thống kê của chuỗi thời gian và phân tích trên chuỗi mã này sẽ giúp dễ dàng trích xuất các thông tin đặc trưng để phân biệt các dạng khác nhau trong chuỗi dữ liệu thời gian gốc.

Cấu trúc của khối trích chọn các đặc trưng dựa trên các chuỗi mã ký tự hóa 503 được minh họa trong hình 7 gồm có các khối con là: khối ký tự hóa dữ liệu chuỗi thời gian 701, khối ký tự hóa chuỗi sai phân bậc một của dữ liệu chuỗi thời gian 702, khối tính toán đặc trưng dựa trên chuỗi mã của dữ liệu chuỗi thời gian 703 và khối tính toán đặc trưng dựa trên chuỗi mã của chuỗi sai phân bậc một 704.

Đầu vào của khối trích chọn các đặc trưng dựa trên chuỗi mã ký tự hóa 503 là dữ liệu chuỗi thời gian đã qua khối tiền xử lý 201.

- Dữ liệu chuỗi thời gian sau tiền xử lý được đưa qua khối ký tự hóa dữ liệu chuỗi thời gian 701 thực hiện thuật toán để sinh chuỗi mã của dữ liệu chuỗi thời gian $\{c_p[n]\}$. Chuỗi mã này là đầu vào của khối tính toán đặc trưng chỉ số đo mức độ hỗn loạn entropý (entropy) chuẩn hóa của đồ thị tần suất của chuỗi mã $\{c_p[n]\}$ 703, khối này sẽ thực hiện tính toán đặc trưng của dữ liệu chuỗi thời gian theo công thức (CT5):

$$f_4 = -\frac{1}{\log_2 N_q} \sum_{i=1}^{N_q} q_i \log_2 q_i \quad (\text{CT5})$$

trong đó, đồ thị tần suất của chuỗi mã là một vectơ xác suất $(q_1, \dots, q_{N_q})$ với N_q là số giá trị khác nhau trong chuỗi mã và q_i là xác suất xuất hiện của giá trị thứ i . Đặc trưng f_4 giúp phân biệt dạng biến thiên ngẫu nhiên (do dữ liệu chuỗi thời gian biến thiên ngẫu nhiên nên entropy của nó là lớn nhất) so với các dạng biến thiên còn lại.

- Dữ liệu chuỗi thời gian này đồng thời cũng được đưa qua khối ký tự hóa chuỗi sai phân bậc một của dữ liệu chuỗi thời gian 702. Đầu tiên, khối này sẽ tạo ra chuỗi sai phân bậc một của dữ liệu chuỗi thời gian $\{d[n]\}$ được định nghĩa là $d[n] = p'[n+1] - p'[n]; n = 0, \dots, N-2$). Sau đó, dựa trên chuỗi $\{d[n]\}$, khối này thực hiện thuật toán để tạo ra chuỗi mã cho chuỗi sai phân bậc một $\{c_d[n]\}$. Chuỗi mã được tạo ra là đầu vào của khối tính toán hai đặc trưng dựa trên chuỗi mã của chuỗi sai phân bậc một dữ liệu chuỗi thời gian 704. Khối này tính toán hai đặc trưng là độ nhọn (kurtosis) và mômen bậc hai chuẩn hóa (normalized second-order moment) của chuỗi mã $\{c_d[n]\}$ theo hai công thức (CT6) và (CT7) tương ứng:

$$f_5 = \frac{\frac{1}{N-3} \sum_{i=0}^{N-4} (c_d[i] - \mu_{c_d})^4}{\left(\frac{1}{N-3} \sum_{i=0}^{N-4} (c_d[i] - \mu_{c_d})^2\right)^2} - 3 \quad (\text{CT6})$$

$$f_6 = \frac{1}{N-3} \sum_{i=0}^{N-4} \left(\frac{c_d[i]}{26}\right)^2 \quad (\text{CT7})$$

trong đó, μ_{c_d} là giá trị trung bình của chuỗi mã $\{c_d[n]\}$.

Các công thức (CT5), (CT6) và (CT7) được tham khảo dựa trên nội dung bài báo “Phát hiện dạng chu kỳ lặp xung sử dụng (kỹ thuật) ký tự hóa” (“Pulse repetition interval modulation recognition using symbolization”) của K. H. Song, D.W. Lee, J. W. Han và B. K. Park công bố trên kỷ yếu “Hội nghị quốc tế về tính toán ảnh số: các kỹ thuật và ứng dụng” (“International Conference on Digital Image Computing: Techniques and Applications”) xuất bản năm 2010 (từ trang 540 đến trang 545). Đặc trưng f_5 là chỉ số để đánh giá độ nhọn của phân bố chuỗi mã. Do dạng biến thiên theo chu kỳ có độ nhọn phân bố chuỗi mã tương đối nhỏ, đặc trưng này giúp phân biệt dạng biến thiên này so với các dạng biến thiên còn lại. Đặc trưng f_6 giúp phân biệt các dạng biến thiên tuyến tính tăng, tuyến tính giảm và nhóm (giá trị mômen chuẩn hóa biến đổi xung quanh giá trị 1, 0 và 0.25).

- Khối phân loại dựa trên mạng nơron

Tham chiếu hình 2 và hình 8, khối phân loại 203 là một mạng nơron gồm 4 lớp đã huấn luyện. Mạng nơron là một mô hình tính toán được xây dựng dựa trên ý tưởng các mạng nơron sinh học gồm có một nhóm các nơron nhân tạo (nút) nối với nhau và xử lý thông tin bằng cách truyền theo các kết nối và tính giá trị mới tại các nút. Mạng nơron được sử dụng là mạng có cấu trúc truyền thẳng gồm 4 lớp: lớp đầu vào có kích thước 6 (đầu vào của mạng là 6 đặc trưng đã được trích chọn ở khối trích chọn các đặc trưng 202), lớp ẩn (hidden layer) là một lớp kết nối đầy đủ gồm có 6 nơron, lớp cực đại mềm (softmax) thực hiện hàm cực đại mềm (softmax) có kích thước 7 và lớp đầu ra có kích thước 7 (tương ứng với số lượng dạng biến thiên dữ liệu chuỗi thời gian cần phân loại). Tổng số tham số của mạng là 91.

Khối phân loại 203 thực hiện các phép toán nhân/cộng ma trận giữa các tham số của mạng với bộ 6 đặc trưng đầu vào. Đầu ra của lớp đầu ra là một vectơ có kích thước là 7, trong đó giá trị ứng với dạng biến thiên được xác định là lớn nhất. Do đó đầu ra của khối phân loại 203 là dạng biến thiên của dữ liệu chuỗi thời gian đầu vào (một trong 7 dạng biến thiên minh họa ở hình 1a đến hình 1g: cố định, tuyến tính tăng, tuyến tính giảm, ngẫu nhiên, so le, bậc thang và chu kỳ).

- Khối ước lượng các tham số

Tham chiếu hình 2, khối ước lượng các tham số dạng biến thiên dữ liệu chuỗi thời gian 204 có chức năng phân tích và ước lượng các tham số ứng với dạng biến thiên đã được xác định ở khối phân loại 203 (số mức, giá trị trung bình, giá trị lớn nhất, giá trị nhỏ nhất và độ lệch chuẩn của dữ liệu chuỗi thời gian đầu vào).

Tham chiếu hình 9, cấu trúc của khối này gồm có ba khối con là: khối phân loại nhóm dạng biến thiên 901, khối ước lượng tham số cho dạng biến thiên một mức 902 và khối ước lượng tham số cho dạng biến thiên nhiều mức 903.

Đầu vào của khối ước lượng các tham số là dữ liệu chuỗi thời gian sau tiền xử lý và dạng biến thiên đã được xác định ở khối phân loại 203. Khối phân loại nhóm dạng biến thiên 901 dựa trên dạng biến thiên của dữ liệu chuỗi thời gian để phân loại dạng biến thiên này thuộc nhóm dạng biến thiên một mức hay thuộc nhóm dạng biến thiên nhiều mức giá trị rời rạc để đưa vào các khối ước lượng tham số tương ứng:

- Đối với các dạng biến thiên thuộc nhóm dạng biến thiên một mức, dữ liệu chuỗi thời gian sẽ được đưa vào khối ước lượng tham số cho dạng biến thiên một mức 902. Khối này thực hiện các phép toán học để tính giá trị trung bình, giá trị lớn nhất, giá trị nhỏ nhất và giá trị độ lệch chuẩn của dữ liệu chuỗi thời gian đầu vào.
- Đối với các dạng biến thiên thuộc nhóm dạng biến thiên nhiều mức, dữ liệu chuỗi thời gian sẽ được đưa vào khối ước lượng tham số cho dạng biến thiên nhiều mức 903. Khối này gồm có hai khối con được minh họa trong hình 10 gồm: khối ước lượng số mức giá trị trong dữ liệu 10001 và khối ước lượng tham số ứng với mỗi mức giá trị 10002. Khối ước lượng số mức trong dữ liệu 10001 xây dựng đồ thị tần suất cho dữ liệu chuỗi thời gian. Do dữ liệu chuỗi thời gian biến thiên lặp lại theo từng mức, số đỉnh trong đồ thị tần suất sẽ tương ứng với các mức giá trị trong dữ liệu chuỗi thời gian. Tương ứng với mỗi mức đã xác định trong khối ước lượng số mức giá trị trong dữ liệu 10001, khối ước lượng tham số ứng với mỗi mức giá trị 10002 thực hiện các phép toán học để tính giá trị trung bình, giá trị lớn nhất, giá trị nhỏ nhất và giá trị độ lệch chuẩn của dữ liệu chuỗi thời gian đầu vào tại từng mức giá trị trong dữ liệu chuỗi thời gian.

Kết quả đầu ra của khối ước lượng các tham số 204, cũng là kết quả đầu ra của toàn hệ thống là dạng biến thiên của dữ liệu chuỗi thời gian, số mức trong dữ liệu chuỗi thời gian và các tham số ước lượng tương ứng với số mức đó trong dữ liệu chuỗi thời gian (giá trị trung bình, giá trị lớn nhất, giá trị nhỏ nhất và giá trị độ lệch chuẩn).

Để đạt được mục đích thứ hai của sáng chế là đề xuất phương pháp tự động phân loại và ước lượng các dạng biến thiên của dữ liệu chuỗi thời gian, phương pháp này bao gồm các bước như sau: bước 1: thực hiện lọc trung vị đối với dữ liệu chuỗi thời gian đầu vào; bước 2: trích chọn các đặc trưng của các dạng biến thiên dữ liệu chuỗi thời gian; bước 3: xác định dạng biến thiên dữ liệu chuỗi thời gian; bước 4: ước lượng các đặc trưng ứng với dạng biến thiên dữ liệu chuỗi thời gian. Phương pháp tự động phân loại, ước lượng dạng biến thiên dữ liệu chuỗi thời gian được đề xuất dựa trên sử dụng hệ thống nói trên cụ thể như sau:

- Bước 1: thực hiện lọc trung vị đối với dữ liệu chuỗi thời gian đầu vào

Dữ liệu chuỗi thời gian đầu vào được thực hiện lọc trung vị sử dụng bộ lọc trung vị trong khối tiền xử lý 201 để giảm bớt ảnh hưởng của nhiễu lên dữ liệu chuỗi thời gian đầu vào. Bước này được thực hiện thông qua máy tính được lập trình sẵn.

- Bước 2: trích chọn các đặc trưng của các dạng biến thiên dữ liệu chuỗi thời gian

Dữ liệu chuỗi thời gian sau khi được lọc trung vị ở khối tiền xử lý 201 được đưa vào khối trích chọn các đặc trưng 202 để trích xuất ra 6 đặc trưng, cụ thể như sau:

- Dữ liệu chuỗi thời gian được đưa vào khối con trích chọn các đặc trưng dựa trên hàm tự tương quan 501. Hàm tự tương quan của dữ liệu chuỗi thời gian được tính toán sử dụng khối con tính toán hàm tự tương quan 601 theo công thức (CT1). Hai đặc trưng là tỷ lệ giữa độ biến thiên của chuỗi thời gian với giá trị trung bình và tỷ lệ giữa giá trị hàm tự tương quan của dữ liệu chuỗi thời gian

với trể 0 mẫu và trể 1 mẫu trích xuất theo công thức (CT2) và (CT3) được tính toán bởi khối tính toán 2 đặc trưng dựa trên hàm tự tương quan 602.

- Dữ liệu chuỗi thời gian được đưa vào khối con trích chọn đặc trưng dựa trên số nhóm 502. Khối này trích xuất đặc trưng số lượng nhóm khác nhau trong dữ liệu chuỗi thời gian sử dụng thuật toán ước lượng số nhóm cụ thể như sau:
 - o Thiết lập một ngưỡng để chia nhóm T_c , khởi tạo số lượng nhóm bằng 0;
 - o Sắp xếp các giá trị trong dữ liệu chuỗi thời gian $\{p'[n]\}$ theo thứ tự tăng dần để thu được chuỗi $\{p'_{sx}[n]\}$;
 - o Lần lượt duyệt qua các phần tử trong chuỗi sắp xếp $\{p'_{sx}[n]\}$ nếu $p'_{sx}[n+1] - p'_{sx}[n] > T_c$ thì hai phần tử này được xếp vào hai nhóm khác nhau, tức là số nhóm tăng thêm 1.
- Dữ liệu chuỗi thời gian được đưa vào khối con trích chọn các đặc trưng dựa trên các chuỗi mã ký tự hóa 503. Khối ký tự hóa dữ liệu chuỗi thời gian 701 thực hiện thuật toán để sinh ra chuỗi mã của dữ liệu chuỗi thời gian $\{c_p[n]\}$ theo nguyên lý sau:
 - o Ánh xạ dữ liệu chuỗi thời gian sang chuỗi ký tự $\{s_p[n]\}$ theo nguyên tắc

$$s_p[n] = \begin{cases} 0, & p'[n] < \mu_p \\ 1, & p'[n] \geq \mu_p \end{cases}, \text{ trong đó } \mu_p \text{ là giá trị trung bình của } \{p'[n]\};$$
 - o Chuỗi mã của dữ liệu chuỗi thời gian được tạo ra dựa trên chuỗi ký tự $\{s_p[n]\}$ theo nguyên tắc $c_p[n] = s_p[n] \times 2^2 + s_p[n+1] \times 2^1 + s_p[n+2] \times 2^0$.

Khối ký tự hóa dữ liệu chuỗi sai phân bậc một của dữ liệu chuỗi thời gian 702 tạo ra chuỗi sai phân bậc một của dữ liệu chuỗi thời gian $\{d[n]\}$. Sau đó, dựa trên chuỗi $\{d[n]\}$, khối này thực hiện thuật toán để tạo ra chuỗi mã cho chuỗi sai phân bậc một $\{c_d[n]\}$ theo nguyên lý như sau:

- o Ánh xạ chuỗi sai phân bậc một của dữ liệu chuỗi thời gian $\{d[n]\}$ sang chuỗi ký tự $\{s_d[n]\}$ theo nguyên tắc $s_d[n] = \begin{cases} 0, & d[n] < -\varepsilon \\ 1, & |d[n]| \leq \varepsilon \\ 2, & d[n] > \varepsilon \end{cases}$ trong đó, ε là một ngưỡng xác định;

- Chuỗi mã của chuỗi sai phân bậc một của dữ liệu chuỗi thời gian được tạo ra dựa trên chuỗi ký tự $\{s_d[n]\}$ theo nguyên tắc $c_d[n] = s_d[n] \times 3^2 + s_d[n+1] \times 3^1 + s_d[n+2] \times 3^0$.
- Hai chuỗi mã tạo ra là đầu vào cho các khối tính toán 703 và 704 để trích xuất 3 đặc trưng là chỉ số đo mức độ hỗn loạn entropi (entropy) chuẩn hóa của đồ thị tần suất của chuỗi mã, độ nhọn (kurtosis) của phân bố chuỗi mã của dữ liệu chuỗi thời gian và mômen bậc hai chuẩn hóa (normalized second-order moment) của chuỗi mã của chuỗi sai phân bậc một theo các công thức (CT5), (CT6) và (CT7).
- Bước 3: xác định dạng biến thiên dữ liệu chuỗi thời gian

6 đặc trưng đã được trích xuất ở bước 2 là đầu vào cho khối phân loại 203. Bộ 6 vectơ đặc trưng này được thực hiện các phép toán nhân/cộng ma trận với bộ hệ số của mạng nơron truyền thẳng đã được huấn luyện để thu được kết quả là một vectơ có kích thước là 7, trong đó giá trị ứng với dạng biến thiên được xác định là lớn nhất. Kết quả đầu ra sau khi tính toán ở khối phân loại 203 là dạng biến thiên của dữ liệu chuỗi thời gian được xác định.

- Bước 4: ước lượng các đặc trưng ứng với dạng biến thiên dữ liệu chuỗi thời gian

Sau khi đã xác định được dạng biến thiên của dữ liệu chuỗi thời gian ở bước 3, các tham số đặc trưng của dữ liệu chuỗi thời gian được ước lượng sử dụng khối ước lượng các tham số 204. Quá trình xử lý ở bước 4 được minh họa ở hình 11, cụ thể như sau:

Đầu tiên, dạng biến thiên của dữ liệu chuỗi thời gian sẽ được phân vào nhóm dạng biến thiên một mức hay nhiều mức sử dụng khối phân nhóm dạng biến thiên 901. Khối này kiểm tra dạng biến thiên của dữ liệu chuỗi thời gian đã xác định ở bước 3 nếu là dạng biến thiên so le (staggered) hoặc dạng biến thiên bậc thang (dwell&switch) thì phân loại vào nhóm dạng biến thiên nhiều mức rời rạc, nếu là các dạng biến thiên còn lại (cố định, ngẫu nhiên, tuyến tính tăng, tuyến tính giảm và theo chu kỳ) thì phân loại vào nhóm dạng biến thiên một mức.

- + Đối với nhóm dạng biến thiên một mức, dữ liệu chuỗi thời gian sau tiền xử lý được đưa vào khối ước lượng tham số cho dạng biến thiên một mức 902, khối này thực hiện các công thức (CT8 - CT11) để tính giá trị trung bình, giá trị lớn nhất, giá trị nhỏ nhất và giá trị độ lệch chuẩn của dữ liệu chuỗi thời gian $\{p'[n]\}$:

$$\blacksquare \text{ Giá trị trung bình: } Mean = \frac{\sum_{n=0}^{N-1} p'[n]}{N} \quad (CT8)$$

$$\blacksquare \text{ Giá trị lớn nhất: } Max = \max(p'[n]) \quad (CT9)$$

$$\blacksquare \text{ Giá trị nhỏ nhất: } Min = \min(p'[n]) \quad (CT10)$$

$$\blacksquare \text{ Độ lệch chuẩn: } Std = \sqrt{\frac{\sum_{n=0}^{N-1} (p'[n] - Mean)^2}{N}} \quad (CT11)$$

với $n = 0, \dots, N - 1$;

- + Đối với nhóm dạng biến thiên nhiều mức, dữ liệu chuỗi thời gian sau tiền xử lý được đưa vào khối ước lượng tham số cho dạng biến thiên nhiều mức 903.

Đồ thị tần suất (histogram) của dữ liệu chuỗi thời gian được tính toán trong khối ước lượng số mức trong dữ liệu 10001 như sau:

- chuyển dữ liệu chuỗi thời gian đầu vào sang dạng khoảng, gọi là các ngăn (bin); mỗi ngăn đại diện cho một dải giá trị trong dữ liệu chuỗi thời gian đầu vào; kích cỡ của mỗi ngăn là cố định.
- Đối với mỗi ngăn, đếm số điểm dữ liệu trong dữ liệu chuỗi thời gian thỏa mãn điều kiện của ngăn (lớn hơn giá trị nhỏ nhất của ngăn và nhỏ hơn hoặc bằng giá trị lớn nhất của ngăn).
- Đồ thị tần suất có trục hoành là giá trị dữ liệu chuỗi thời gian chia theo các ngăn và trục tung là số điểm trong dữ liệu chuỗi thời gian tương ứng với các ngăn (tần suất xuất hiện).

Sau khi xây dựng đồ thị tần suất, số mức giá trị của dạng biến thiên dữ liệu chuỗi thời gian được xác định bằng số đỉnh trong đồ thị tần suất. Các đỉnh trong đồ thị tần suất được xác định là các giá trị tại các ngăn lớn hơn một ngưỡng T xác định trước. Hình 12a và hình 12b minh họa đồ thị tần suất và các đỉnh trong đồ thị tần suất này với trường hợp dạng biến thiên bậc thang

với số mức giá trị khác nhau bằng 3. Tương ứng với mỗi mức giá trị trong dữ liệu chuỗi thời gian l_i (một đỉnh trong đồ thị tần suất), các điểm dữ liệu tương ứng trong dữ liệu chuỗi thời gian được đưa vào khối ước lượng tham số ứng với mỗi mức 1002. Khối này thực hiện các công thức toán học (CT12-CT15) để ước lượng giá trị trung bình, giá trị lớn nhất, giá trị nhỏ nhất và giá trị độ lệch chuẩn của dữ liệu chuỗi thời gian ứng với một mức giá trị:

- Giá trị trung bình: $Mean(l_i) = \frac{\sum_{n=0}^{N_l-1} p_l[n]}{N}$ (CT12)

- Giá trị lớn nhất: $Max(l_i) = \max(p_l[n])$ (CT13)

- Giá trị nhỏ nhất: $Min(l_i) = \min(p_l[n])$ (CT14)

- Độ lệch chuẩn: $Std(l_i) = \sqrt{\frac{\sum_{n=0}^{N_l-1} (p_l[n] - Mean(l_i))^2}{N_l}}$ (CT15)

trong đó, $\{p_l[n]\}, n = 0, \dots, N_l - 1$ là dữ liệu chuỗi thời gian trong ngăn của đồ thị tần suất ứng với mức giá trị l_i ; N_l là kích thước của chuỗi $\{p_l[n]\}$.

Đầu ra của khối ước lượng các tham số 204 tại bước 4, cũng là đầu ra của hệ thống là dạng biến thiên của dữ liệu chuỗi thời gian, số mức giá trị trong dữ liệu chuỗi thời gian và các tham số ước lượng của dữ liệu chuỗi thời gian (giá trị trung bình, giá trị lớn nhất, giá trị nhỏ nhất và giá trị độ lệch chuẩn).

Yêu cầu bảo hộ

1. Hệ thống tự động phân loại, ước lượng các dạng biến thiên của dữ liệu chuỗi thời gian bao gồm các khối:

khối tiền xử lý; khối này có chức năng giảm ảnh hưởng của nhiễu đối với dữ liệu chuỗi thời gian đầu vào, khối tiền xử lý là một bộ lọc trung vị có kích thước của sổ là K mẫu, đầu ra của khối là dữ liệu chuỗi thời gian đã được lọc bỏ nhiễu;

khối trích chọn các đặc trưng; khối này bao gồm các khối con có chức năng phân tích và trích xuất 6 đặc trưng cho dạng biến thiên dữ liệu chuỗi thời gian từ dữ liệu chuỗi thời gian đầu vào sau tiền xử lý, 6 đặc trưng này chính là đầu ra sau khi dữ liệu chuỗi thời gian được xử lý qua khối trích chọn các đặc trưng, bao gồm:

- tỷ lệ giữa độ biến thiên của dữ liệu chuỗi thời gian với giá trị trung bình;
- tỷ lệ giữa giá trị hàm tự tương quan của dữ liệu chuỗi thời gian với trễ 0 mẫu và trễ 1 mẫu;
- số lượng nhóm khác nhau trong dữ liệu chuỗi thời gian;
- chỉ số đo mức độ hỗn loạn entropy (entropy) chuẩn hóa của đồ thị tần suất của chuỗi mã hóa dữ liệu chuỗi thời gian;
- độ nhọn (kurtosis) của phân bố của chuỗi mã của dữ liệu chuỗi thời gian;
- mômen bậc hai chuẩn hóa (normalized second-order moment) của chuỗi mã của chuỗi sai phân bậc một dữ liệu chuỗi thời gian;

khối phân loại gồm một mạng nơron truyền thẳng đã huấn luyện gồm 4 lớp bao gồm: lớp đầu vào có kích thước 6 (là 6 đặc trưng đã được trích chọn ở khối trích chọn các đặc trưng), lớp ẩn (hidden layer) là một lớp kết nối đầy đủ gồm 6 nơron, lớp cực đại mềm (softmax) thực hiện hàm softmax có kích thước 7 và lớp đầu ra có kích thước 7 (tương ứng với số lượng dạng biến thiên dữ liệu chuỗi thời gian cần phân loại), tổng số tham số của mạng là 91; khối phân loại thực hiện các phép toán nhân/cộng ma trận

giữa đầu và của mạng là bộ 6 đặc trưng; đầu ra là một véctơ có kích thước là 7, trong đó giá trị ứng với dạng biến thiên được xác định là lớn nhất, đây cũng chính là dạng biến thiên của dữ liệu chuỗi thời gian đầu vào;

khối ước lượng các tham số; đầu vào của khối ước lượng các tham số là dữ liệu chuỗi thời gian sau tiền xử lý và dạng biến thiên đã được xác định ở khối phân loại; khối phân loại nhóm dạng biến thiên dựa trên dạng biến thiên của dữ liệu chuỗi thời gian để phân loại dạng biến thiên này thuộc nhóm dạng biến thiên một mức hay thuộc nhóm dạng biến thiên nhiều mức giá trị rời rạc để đưa vào các khối ước lượng tham số tương ứng; kết quả đầu ra của khối ước lượng các tham số, cũng là kết quả đầu ra của toàn hệ thống là dạng biến thiên của dữ liệu chuỗi thời gian, số mức trong dữ liệu chuỗi thời gian và các tham số ước lượng tương ứng với số mức đó trong dữ liệu chuỗi thời gian (giá trị trung bình, giá trị lớn nhất, giá trị nhỏ nhất và giá trị độ lệch chuẩn).

2. Hệ thống tự động phân loại, ước lượng các dạng biến thiên của dữ liệu chuỗi thời gian theo điểm 1 có dữ liệu chuỗi thời gian là chuỗi các điểm dữ liệu rời rạc là kết quả của việc đo lường một đại lượng trong những khoảng thời gian liên tiếp nhau với tần suất thời gian thống nhất; dữ liệu chuỗi thời gian này có độ dài thay đổi giữa các lần phân tích khác nhau và thường biến thiên theo một dạng xác định.

3. Hệ thống tự động phân loại, ước lượng các dạng biến thiên của dữ liệu chuỗi thời gian theo điểm 1, trong đó, khối trích chọn các đặc trưng bao gồm các khối con:

khối trích chọn các đặc trưng dựa trên hàm tự tương quan; khối này gồm hai khối con là: khối tính toán hàm tự tương quan và khối tính toán 2 đặc trưng dựa trên hàm tự tương quan; đầu vào của khối là dữ liệu chuỗi thời gian đã qua khối tiền xử lý; đầu ra của khối là hai đặc trưng được trích chọn cho dữ liệu chuỗi thời gian: đặc trưng đầu tiên giúp phân biệt dạng dữ liệu chuỗi thời gian cố định (biến thiên rất nhỏ) với các dạng biến thiên khác, đặc trưng thứ hai giúp phân biệt dạng biến thiên dữ liệu chuỗi thời gian so le (khác biệt giữa các mẫu liên tiếp nhau thường lớn) với các dạng biến thiên khác;

khối trích chọn đặc trưng ước lượng số nhóm trong dữ liệu chuỗi thời gian; khối này giúp phân biệt các dạng biến thiên liên tục giữa các mẫu với các dạng biến

thiên gồm nhiều mức giá trị rời rạc (dạng biến thiên so le và dạng biến thiên bậc thang);

khối trích chọn các đặc trưng dựa trên chuỗi mã ký tự hóa: đầu vào của khối trích chọn các đặc trưng dựa trên chuỗi mã ký tự hóa là dữ liệu chuỗi thời gian đã qua khối tiền xử lý, đầu ra của khối là ba đặc trưng: đặc trưng thứ nhất giúp phân biệt dạng biến thiên ngẫu nhiên (do dữ liệu chuỗi thời gian biến thiên ngẫu nhiên nên entropy của nó là lớn nhất) so với các dạng biến thiên còn lại; đặc trưng thứ hai là chỉ số để đánh giá độ nhọn của phân bố chuỗi mã; đặc trưng thứ ba giúp phân biệt các dạng biến thiên tuyến tính tăng, tuyến tính giảm và nhóm.

4. Phương pháp tự động phân loại, ước lượng các dạng biến thiên của dữ liệu chuỗi thời gian bao gồm các bước:

bước 1: thực hiện lọc trung vị đối với dữ liệu chuỗi thời gian đầu vào, cụ thể: đưa dữ liệu chuỗi thời gian đầu vào qua khối tiền xử lý để thực hiện lọc trung vị để giảm ảnh hưởng của nhiễu đối với dữ liệu này;

bước 2: trích chọn các đặc trưng của các dạng biến thiên dữ liệu chuỗi thời gian, cụ thể: dữ liệu chuỗi thời gian sau khi được lọc trung vị ở khối tiền xử lý được đưa vào khối trích chọn các đặc trưng (được đưa vào 3 khối con trong khối trích chọn các đặc trưng) để trích xuất ra 6 đặc trưng;

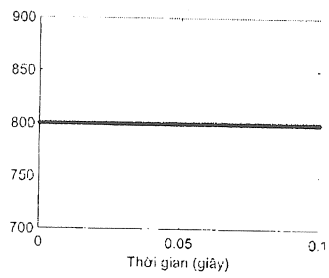
bước 3: xác định dạng biến thiên dữ liệu chuỗi thời gian, cụ thể: 6 đặc trưng đã trích xuất ở bước 2 là đầu vào của khối phân loại; bộ 6 đặc trưng này sẽ thực hiện các phép toán nhân/cộng ma trận với bộ hệ số của một mạng nơron truyền thẳng có kiến trúc như mô tả ở điểm 2 đã được huấn luyện; kết quả đầu ra sau khi thực hiện các phép toán này là giá trị xác định một dạng biến thiên của dữ liệu chuỗi thời gian;

bước 4: ước lượng các đặc trưng ứng với dạng biến thiên dữ liệu chuỗi thời gian, cụ thể: sau khi đã xác định được dạng biến thiên của dữ liệu chuỗi thời gian ở bước 3, các tham số đặc trưng ứng với dạng biến thiên này được ước lượng sử dụng khối ước lượng các tham số, đầu ra của khối ước lượng các tham số tại bước 4, cũng là đầu ra của hệ thống là dạng biến thiên của dữ liệu chuỗi thời gian, số mức giá trị

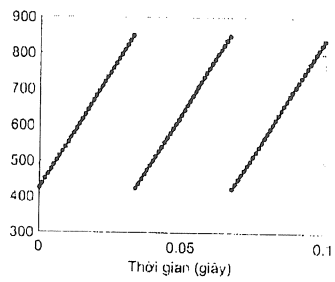
trong dữ liệu chuỗi thời gian và các tham số ước lượng của dữ liệu chuỗi thời gian (giá trị trung bình, giá trị lớn nhất, giá trị nhỏ nhất và giá trị độ lệch chuẩn).

5. Phương pháp tự động phân loại, ước lượng các dạng biến thiên của dữ liệu chuỗi thời gian theo điểm 4 có đầu ra của khối ước lượng tham số sẽ là dạng biến thiên của dữ liệu chuỗi thời gian (đã xác định được ở bước 3) và các tham số bao gồm giá trị trung bình, giá trị lớn nhất, giá trị nhỏ nhất và giá trị độ lệch chuẩn của dữ liệu (được xác định ở bước 4).

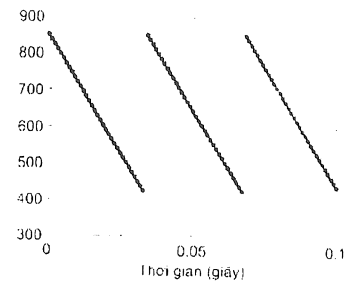
Hình vẽ



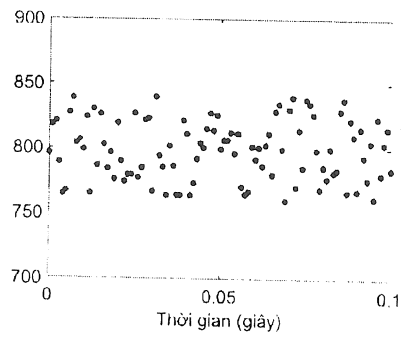
Hình 1a



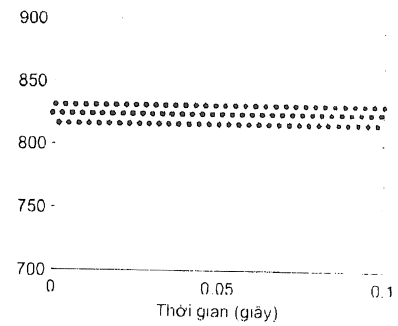
Hình 1b



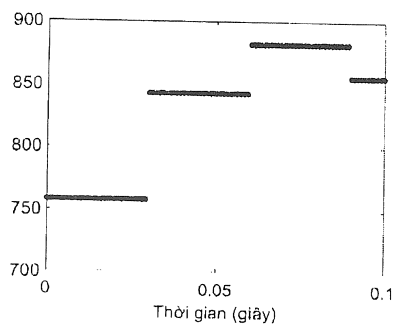
Hình 1c



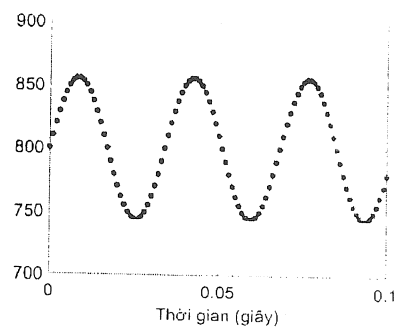
Hình 1d



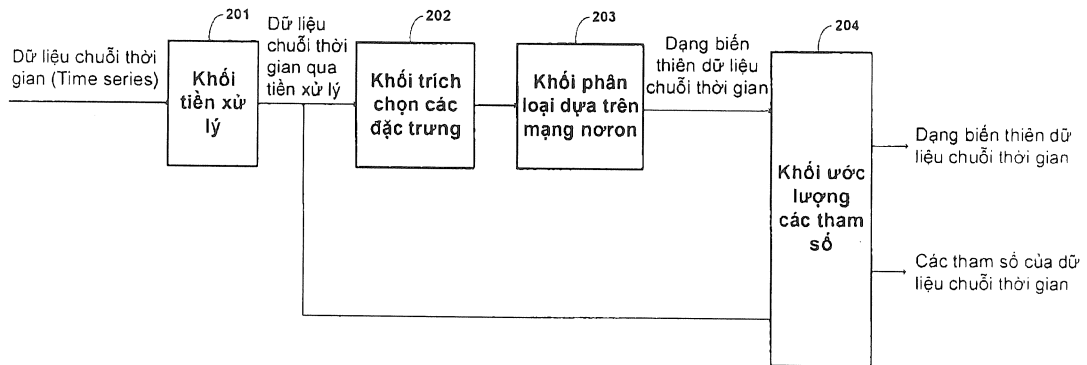
Hình 1e



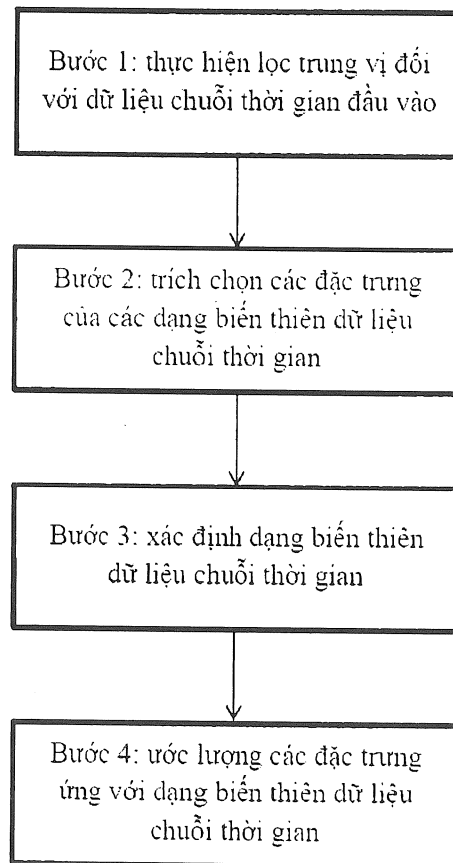
Hình 1f



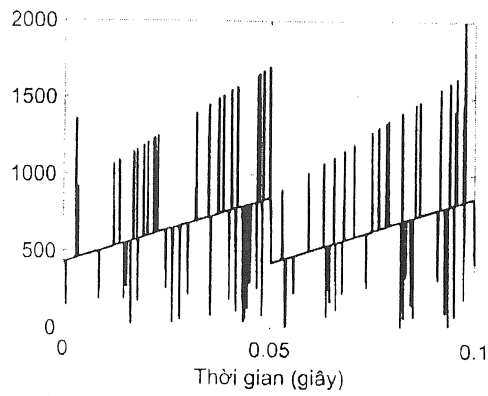
Hình 1g



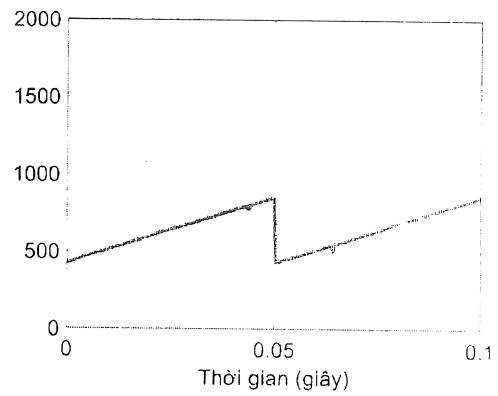
Hình 2



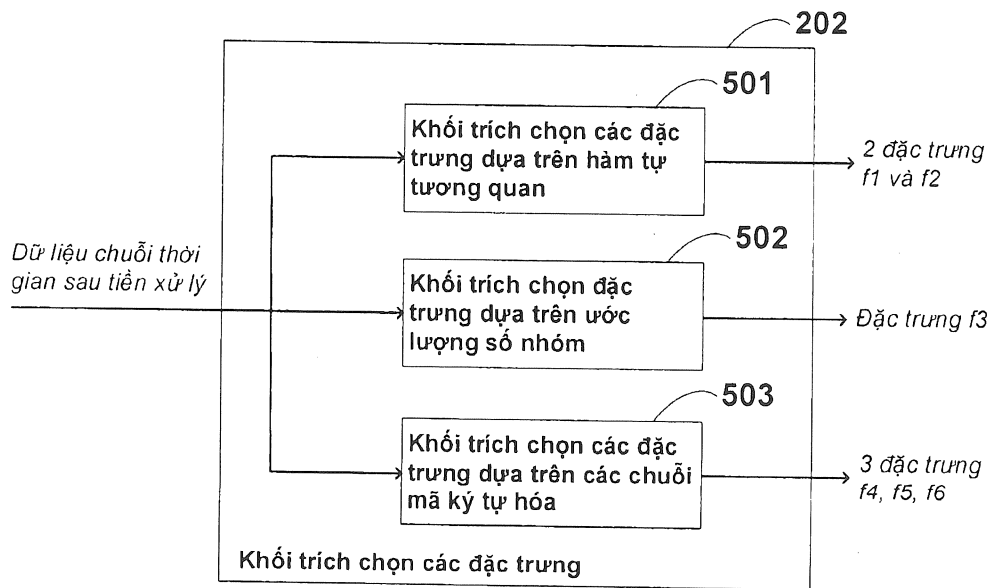
Hình 3



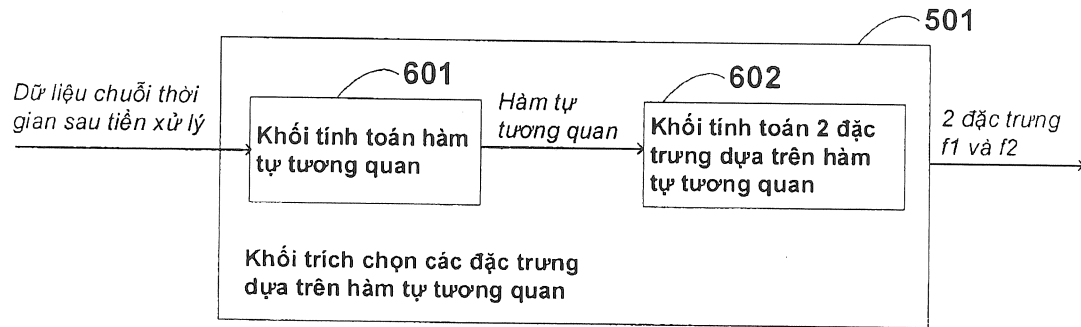
Hình 4a



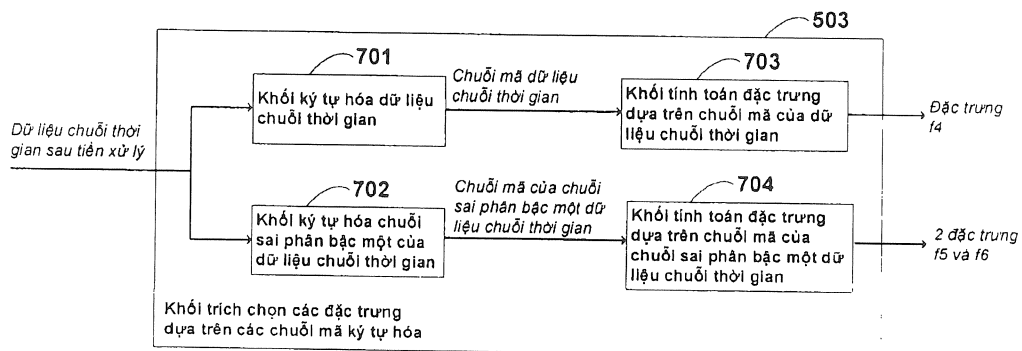
Hình 4b



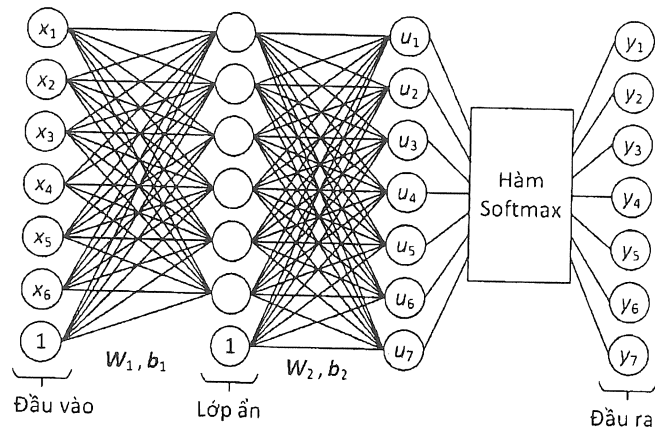
Hình 5



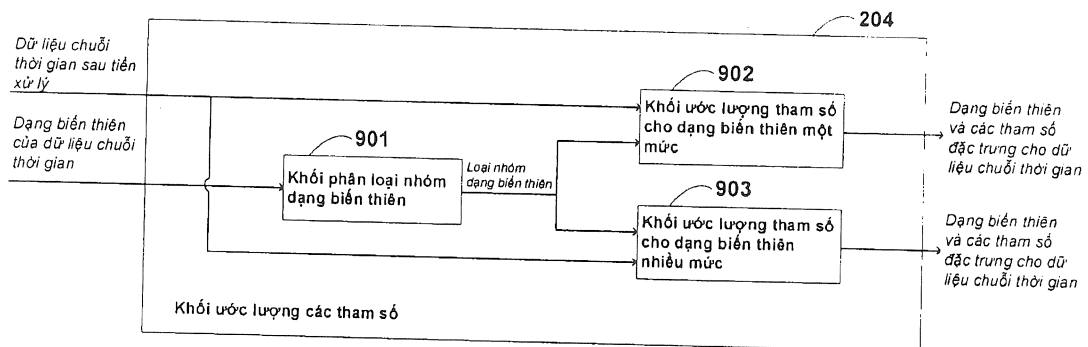
Hình 6



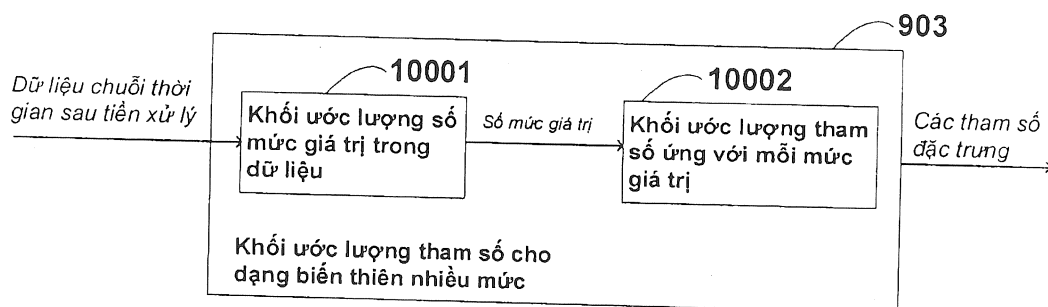
Hình 7



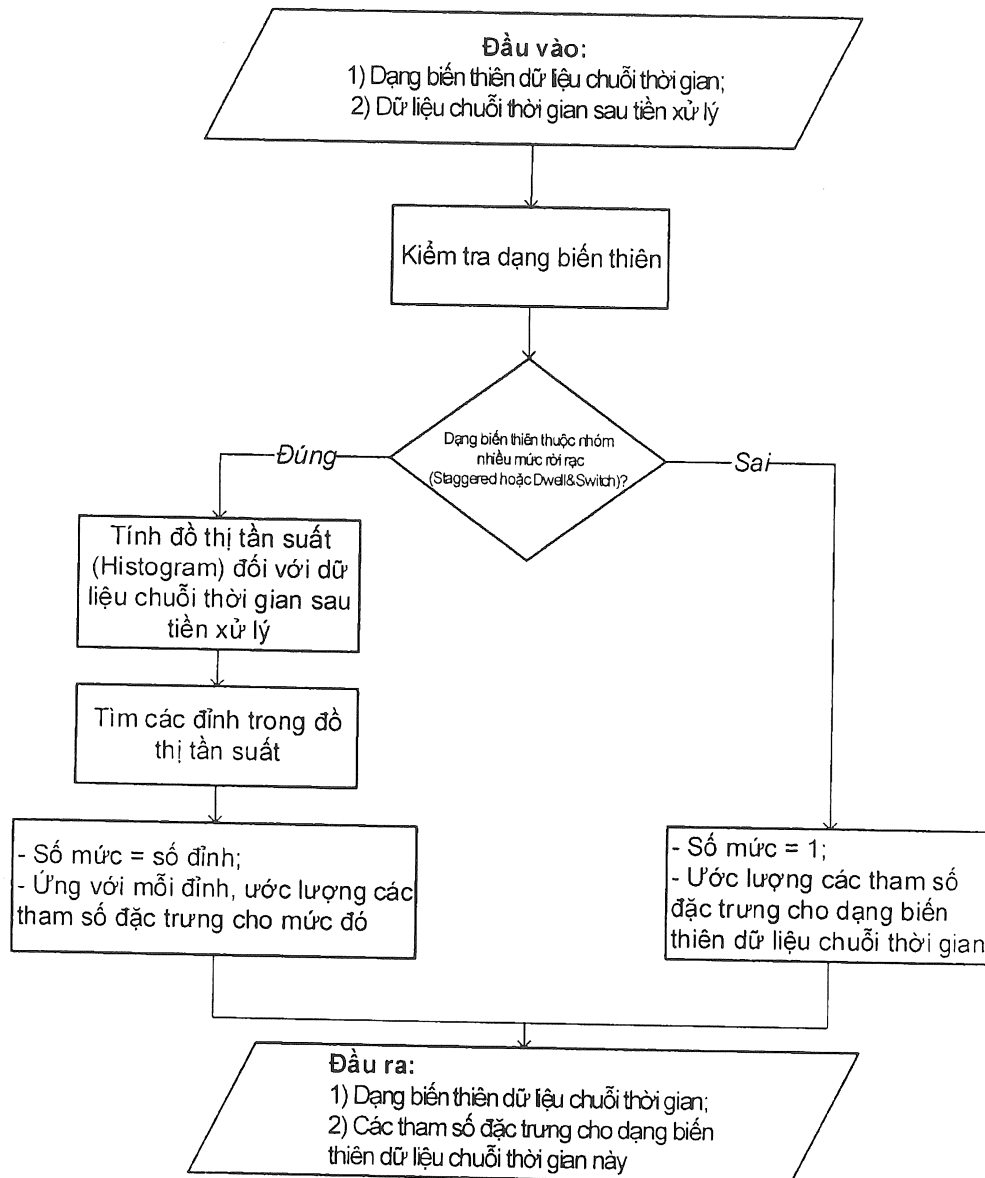
Hình 8



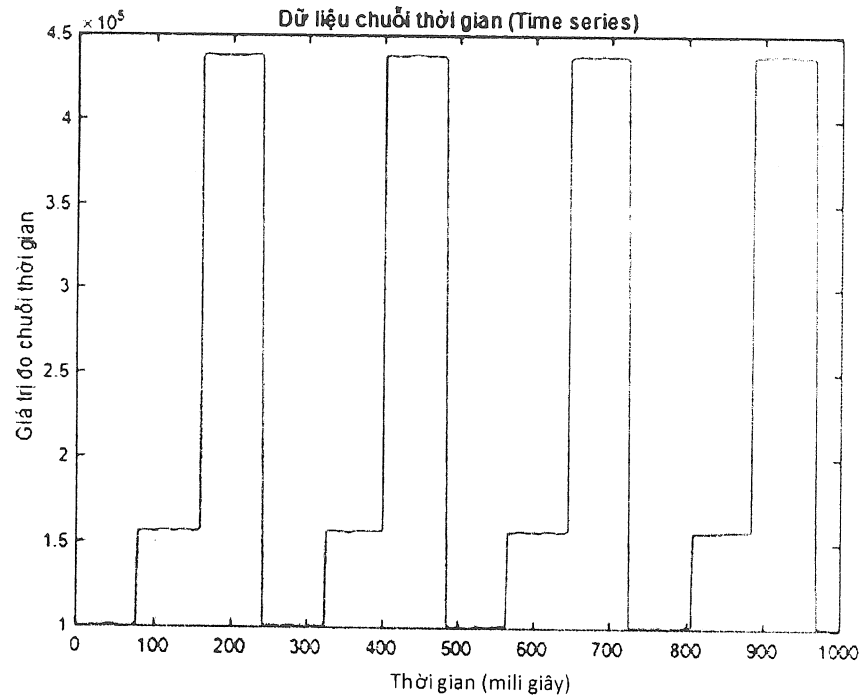
Hình 9



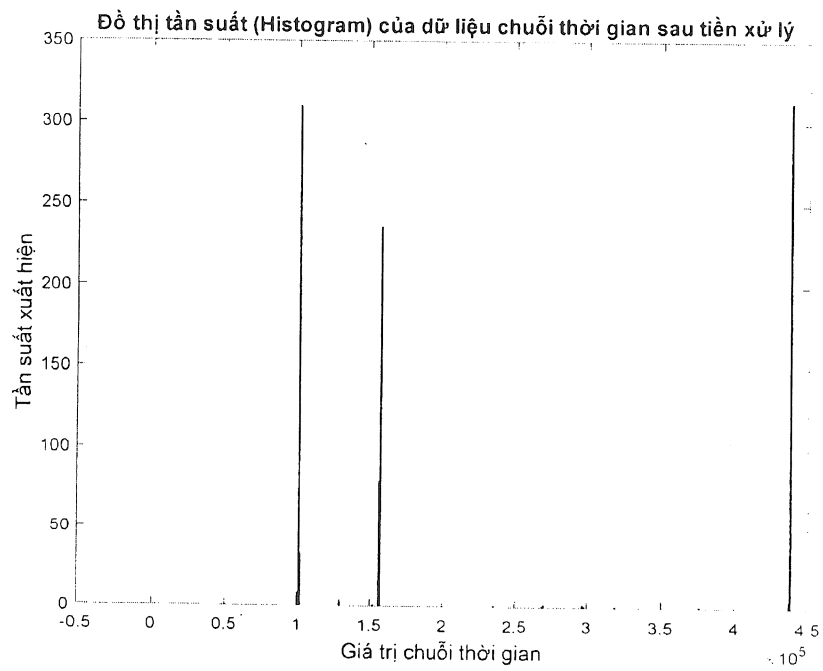
Hình 10



Hình 11



Hình 12a



Hình 12b