

# NHÓM QUAN HỆ MỜ PHỤ THUỘC THỜI GIAN VÀ ỨNG DỤNG TRONG MÔ HÌNH CHUỖI THỜI GIAN MỜ

TS. Nguyễn Công Điều

Khoa Toán-Tin, Đại học Thăng Long

**Tóm tắt:** *Mô hình chuỗi thời gian mờ đang có nhiều ứng dụng trong công tác dự báo, nhất là trong các dự báo kinh tế. Trong những năm gần đây khá nhiều công trình đã được hoàn thành theo hướng nâng cao độ chính xác và giảm khối lượng tính toán trong mô hình chuỗi thời gian mờ như các bài báo của Chen và Hsu, Huarng, Kuo,... Wu. Hầu hết những phương pháp trên đều dựa vào kỹ thuật tạo các nhóm quan hệ logic mờ của Chen [4] để làm giảm khối lượng tính toán khi chỉ cần thực hiện các phép tính số học thay vì các phép tính min-max như trong các mô hình của Song-Chissom. Tuy nhiên các nhóm quan hệ logic mờ này được sử dụng trong mọi quan hệ mờ mà không để ý đến thứ tự xuất hiện và sự lặp lại các thành phần trong vé phôi của mỗi quan hệ logic mờ.*

*Trong bài báo này, chúng tôi đưa ra một cải biên cách xác định nhóm quan hệ mờ phụ thuộc vào thứ tự thời gian. Nhờ có thêm khái niệm nhóm quan hệ mờ và sử dụng mô hình chuỗi thời gian mờ có trọng, chúng tôi đã áp dụng lại được thuật toán hữu hiệu đã được phát triển trước đây. Sử dụng mô hình này chỉ đối với mô hình chuỗi thời gian mờ bậc 1, chúng tôi thu được kết quả dự báo số lượng sinh viên nhập học tốt hơn theo chuẩn MSE so với kết quả của Chen và Yu.*

**Từ khóa:** chuỗi thời gian mờ, mô hình chuỗi thời gian mờ có trọng, một số thuật toán trong mô hình chuỗi thời gian mờ.

## 1. Mở đầu

Chuỗi thời gian mờ và mô hình chuỗi thời gian mờ bậc nhất do Song và Chissom [1]-[3] phát triển từ năm 1993. Sau công trình này, một loạt các bài báo của nhiều tác giả khác nhau tiếp tục dựa trên ý tưởng này để dự báo chuỗi thời gian và ứng dụng trong nhiều lĩnh vực khác nhau như dự báo dân số, tài chính, nhiệt độ, nhu cầu điện, vv... Gần đây có rất nhiều tác giả liên tục cải tiến mô hình chuỗi thời gian mờ để dự báo đạt kết quả chính xác hơn.

Chen [4] đã đưa ra phương pháp mới đơn giản và hữu hiệu hơn so với phương pháp của Song và Chissom bằng cách sử dụng các phép tính số học thay vì các phép tính hợp max-min phức tạp trong xử lý mỗi quan hệ mờ. Phương pháp của Chen chủ yếu dựa trên phương pháp xây dựng nhóm quan hệ logic mờ. Nhiều công trình tiếp theo đã sử dụng cách tiếp cận này để dự báo cho chuỗi thời gian. Huarng [6] đã sử dụng các thông tin có trước trong tính chất của chuỗi thời gian như mức độ tăng giảm để đưa ra mô hình heuristic chuỗi thời gian mờ. Cũng dựa trên tư tưởng này, Dieu [9] đã sử dụng các hàm xác định mối quan hệ heuristic và phân tích hướng pháp triển của chuỗi thời gian để đưa ra các điểm giải mờ.

Trong những năm gần đây, một số tác giả đã sử dụng nhiều kỹ thuật khác nhau để tìm mô hình hữu hiệu cho chuỗi thời gian mờ. Những kỹ thuật trong lý thuyết tính toán mềm, khai phá dữ liệu, mạng nơ ron và các giải thuật tiến hoá đều được đưa vào sử dụng. Một số tác giả sử dụng phương pháp heuristic có tính đến xu hướng như [6,9] hay sử dụng khái niệm tối ưu đám đông như trong các công trình [7], để xây dựng các thuật toán trong mô hình chuỗi thời

gian mờ. Mô hình bậc cao cũng đang được triển khai có hiệu quả bắt đầu từ bài báo của Chen [5]. Ngoài ra, một số tác giả khác đã sử dụng thêm thông tin khác trong chứng khoán để dự báo chính xác hơn các chỉ số chứng khoán. Từ đó nảy sinh ra mô hình chuỗi thời gian mờ 2 nhân tố khi đồng thời với chuỗi thời gian chính còn sử dụng số liệu của các tham số phụ để đưa ra dự báo. Có thể kể ra đây công trình của Lee [8] và một số tác giả khác.

Trong công trình này, chúng tôi đưa ra khái niệm mới là nhóm quan hệ logic mờ phụ thuộc thời gian để nâng cao độ chính xác. Nhận thấy rằng khi xác định nhóm quan hệ mờ, Chen chỉ xác định các các tập mờ có cùng vé trái trong mỗi quan hệ mờ mà không để ý đến lịch sử xuất hiện của từng thành phần của nhóm quan hệ trong vé phải. Trong định nghĩa nhóm quan hệ mờ mới chúng tôi định nghĩa chỉ những phần tử trong vé phải nào xuất hiện trước thời điểm xuất hiện của thành phần vé trái của nhóm quan hệ thì mới tham gia nhóm quan hệ logic mờ. Nhờ có mối quan hệ logic mờ mới này tính toán để giải mờ sẽ đơn giản hơn và cho kết quả tốt hơn so với cách xác định nhóm quan hệ mờ theo Chen. Trong rất nhiều các công trình sau này của các tác giả khác nhau đều dựa trên việc xác định mối quan hệ mờ của Chen để xây dựng giải thuật dự báo. Như vậy với cách cải tiến mới này hy vọng sẽ giúp tăng độ chính xác của dự báo trong các giải thuật khác nhau của mô hình chuỗi thời gian mờ.

Báo cáo này có 5 mục và phần kết luận. Sau phần mở đầu sẽ là phần đưa ra các khái niệm liên quan đến mô hình chuỗi thời gian mờ, đồng thời mô tả các thuật toán cơ bản liên quan đến dự báo thông qua mô hình chuỗi thời gian mờ. Đó là các thuật toán cơ bản của Chen, mô hình có trọng của Yu. Mục 4 đưa ra một cải biên để xác định nhóm quan hệ logic mờ phụ thuộc vào quá trình lịch sử. Mô hình cải biên chuỗi thời gian mờ Mục thứ 5 áp dụng mô hình cải tiến để dự báo số sinh viên nhập học của Đại học Alabama, dự báo chỉ số chứng khoán Đài Loan và xét tính hiệu quả của thuật toán.

## 2. Cơ sở lý thuyết

### 2.1 Một số khái niệm

Trong phần này, chúng ta sẽ sử dụng khái niệm và phương pháp dự báo của chuỗi thời gian mờ được Song và Chissom [1]-[3] phát triển và được Chen [4] cải tiến để xây dựng mô hình dự báo cho chuỗi thời gian.

Một số định nghĩa sau liên quan đến chuỗi thời gian mờ [4].

**Định nghĩa 1:**  $Y(t)$  ( $t = \dots, 0, 1, 2, \dots$ ) là một tập con của  $R^I$ .  $Y(t)$  là tập nền trên đó xác định các tập mờ  $f_i(t)$ .  $F(t)$  là tập chứa các tập  $f_i(t)$  ( $i = 1, 2, \dots$ ). Khi đó ta gọi  $F(t)$  là *chuỗi thời gian mờ* xác định trên tập nền  $Y(t)$ .

**Định nghĩa 2:** Tại các thời điểm  $t$  và  $t-1$  có tồn tại một mối quan hệ mờ giữa  $F(t)$  và  $F(t-1)$  sao cho  $F(t) = F(t-1) * R(t-1, t)$  trong đó  $*$  là ký hiệu của một toán tử xác định trên tập mờ.  $R(t-1, t)$  là *mối quan hệ mờ*. Ta cũng có thể ký hiệu mối quan hệ mờ giữa  $F(t)$  và  $F(t-1)$  bằng  $F(t-1) \rightarrow F(t)$ .

Nếu đặt  $F(t-1) = A_i$  và  $F(t) = A_j$  thì ta ký hiệu *mối quan hệ logic mờ* giữa chúng như sau:  $A_i \rightarrow A_j$ . Viết như thế này có thể hiểu là tập mờ  $A_j$  được suy ra từ  $A_i$ .

**Định nghĩa 3:** Nhóm các mối quan hệ mờ theo Chen.

Các mối quan hệ logic có thể gộp lại thành một nhóm nếu trong ký hiệu trên, cùng một vế trái sẽ có nhiều mối quan hệ tại vế phải. Thí dụ nếu ta có các mối quan hệ:

$$A_i \rightarrow A_k$$

$$A_i \rightarrow A_m$$

thì ta có thể gộp chúng thành *nhóm các mối quan hệ logic* mờ sau:

$$A_i \rightarrow A_k, A_m$$

**Định nghĩa 4:**

Giả sử  $F(t)$  suy ra từ  $F(t-1)$  và  $F(t) = F(t-1) * R(t-1, t)$  cho mọi  $t$ . Nếu  $R(t-1, t)$  không phụ thuộc vào  $t$  thì  $F(t)$  được gọi là chuỗi thời gian mờ dừng, còn ngược lại ta có chuỗi thời gian mờ không dừng.

**Định nghĩa 5:**

Giả sử  $F(t)$  suy đồng thời từ  $F(t-1), F(t-2), \dots, F(t-m)$   $m > 0$  và là chuỗi thời gian mờ dừng. Khi đó mối quan hệ mờ có thể viết được  $F(t-1), F(t-2), \dots, F(t-m) \rightarrow F(t)$  và gọi đó là mô hình dự báo bậc  $m$  của chuỗi thời gian mờ.

**Định nghĩa 6:** Nhóm các mối quan hệ logic mờ theo Yu[10]

Nếu ta có các mối quan hệ :

$$A_i \rightarrow A_k ; , A_i \rightarrow A_m ; A_i \rightarrow A_k$$

Thì nhóm quan hệ mờ theo Yu sẽ được định nghĩa như sau:

$$A_i \rightarrow A_k, A_m, A_k$$

**2.2. Một số Thuật toán trong mô hình chuỗi thời gian mờ**

Thuật toán của Song và Chissom khá phức tạp vì phải tính giá trị max-min trong mối quan hệ mờ. Chen đã có một số cải tiến nên để tính mối quan hệ mờ chỉ cần sử dụng các phép tính số học đơn giản.

Thuật toán của Chen [4] cải tiến thuật toán của Song-Chissom bao gồm một số bước sau:

1. Xác định tập  $U$  bao gồm khoảng giá trị của chuỗi thời gian. Khoảng này xác định từ giá trị nhỏ nhất đến giá trị lớn nhất có thể của chuỗi thời gian.

2. Chia khoảng giá trị và xác định các tập mờ trên tập  $U$ . Vấn đề độ dài của khoảng chưa đặt ra và số lượng khoảng lấy bất kỳ.

3. Mờ hoá các dữ liệu chuỗi thời gian

4. Thiết lập các mối quan hệ logic mờ, nhóm quan hệ logic mờ như Định nghĩa 3.

5. Dự báo và giải mờ. Trong bước dự báo chuỗi thời gian mờ được thực hiện như sau:

- Trường hợp 1: Nếu  $A_j \rightarrow A_i$  và giá trị hàm thuộc của  $A_j$  đạt giá trị *maximum* tại đoạn  $u_i$  và điểm giữa của  $u_i$  là  $m_i$  thì dự báo của chuỗi thời gian tại thời điểm  $i$  là  $m_i$ .

- Trường hợp 2: Nếu ta có các mối quan hệ logic mờ hình thành nhóm quan hệ logic mờ sau:

$$A_i \rightarrow A_{j1}, A_{j2}, \dots, A_{jp}$$

thì giá trị dự báo sẽ là  $A_{i1}, A_{i2}, A_{j1}, \dots, A_{jp}$

Khi đó giải mờ giá trị dự báo sẽ là:

$$\frac{m_{j1} + m_{j2} + \dots + m_{jp}}{p}$$

Trong đó  $m_{j1}, m_{j2}, \dots, m_{jp}$  điểm giữa của các đoạn  $u_i$ .

- *Trường hợp 3*: Nếu vế phải của mỗi quan hệ mờ là trống như trường hợp sau

$$A_i \rightarrow \emptyset$$

thì giá trị dự báo sẽ là  $A_i$  và giải mờ giá trị này sẽ là trung điểm  $m_i$  của đoạn  $u_i$

Yu [16] đã xây dựng mô hình chuỗi thời gian mờ có trọng để xử lý sự lặp lại các tập mờ xuất hiện trong vế phải của nhóm quan hệ mờ. Đối với thứ tự xuất hiện của các tập mờ trong nhóm quan hệ logic mờ ta gán chúng với trọng số khác nhau. Phương pháp này trong đa số các trường hợp cho độ chính xác cao hơn.

### 2.3. Thuật toán cải biên mô hình chuỗi thời gian mờ có trọng

Trước hết ta định nghĩa lại nhóm quan hệ logic mờ. Nhận thấy rằng trong **Định nghĩa 3** nhóm quan hệ mờ không thấy xác định thời gian trong mỗi phần tử của tập mờ  $A_i$ . Chính vì vậy khi nào có nhóm quan hệ logic mờ dạng  $A_i \rightarrow A_{i1}, A_{i2}, \dots, A_{ip}$ , thì ta xử lý giống như khi dự báo giải mờ cho phần tử  $A_i$  không kể phần tử này ứng với giá trị  $t$  khác nhau trong chuỗi thời gian mờ  $F(t)$ . Đáng nhẽ ta phải viết rõ sự tương ứng của phần tử trong chuỗi thời gian mờ là  $F(t-1) = A_i(t)$ . Khi đó trong vế phải của nhóm quan hệ mờ  $A_i \rightarrow A_{i1}, A_{i2}, \dots, A_{ip}$  phải viết lại thành  $A_i(t) \rightarrow A_{i1}(t1), A_{i2}(t2), \dots, A_{ip}(tp)$  chỉ chấp nhận những phần tử nào có thời điểm xuất hiện trước  $t$  mà thôi. Ta sẽ xác định lại nhóm quan hệ logic mờ qua định nghĩa sau.

#### **Định nghĩa 7 (Nhóm quan hệ logic mờ phụ thuộc thời gian)**

Mỗi quan hệ mờ ta đều xác định từ quan hệ  $F(t-1) \rightarrow F(t)$ . Nếu như trên ta đặt  $F(t) = A_i(t)$  và  $F(t-1) = A_j(t-1)$  thì ta có mỗi quan hệ  $A_j(t-1) \rightarrow A_i(t)$ . Nếu tại thời điểm  $t$  ta có các mỗi quan hệ mờ:  $A_j(t-1) \rightarrow A_i(t), A_j(t1-1) \rightarrow A_{i2}(t1), \dots, A_j(tp-1) \rightarrow A_{ip}(tp)$  với các giá trị  $t1, t2, \dots, tp \leq t$  (tức là các mỗi quan hệ mờ trên xảy ra tại các thời điểm trước  $A_j(t-1) \rightarrow A_i(t)$ ) thì ta có thể nhóm các mỗi quan hệ logic mờ thành

$$A_j(t-1) \rightarrow A_i(t), A_{i1}(t1), A_{i2}(t2), \dots, A_{ip}(tp)$$

Và mỗi quan hệ trên được gọi là **nhóm quan hệ logic mờ phụ thuộc thời gian**.

Thực chất cách ghi  $A_j(t)$  vẫn là một tập mờ  $A_j$  đã xác định nhưng chỉ muốn nhấn mạnh tập mờ này xuất hiện tại thời điểm  $t$  mà thôi.

Từ định nghĩa nhóm quan hệ logic này, chúng tôi đưa ra thuật toán giống như thuật toán chuỗi thời gian mờ có trọng của Yu.

1. Xác định tập nền. Tập nền  $U$  được xác định như sau: lấy giá trị lớn nhất  $f_{\max}$  và nhỏ nhất  $f_{\min}$  của chuỗi thời gian và  $U = [f_{\min} - f_1, f_{\max} + f_2]$  trong đó  $f_1, f_2$  là những giá trị dương nào đó. Chia đoạn  $U$  thành  $m$  khoảng con bằng nhau  $u_1, u_2, \dots, u_m$ .

2. Xây dựng các tập mờ  $A_i$  tương ứng với các khoảng con như trong bước 2 và sử dụng các hàm thuộc tam giác cho mỗi khoảng con của phép chia và mờ hoá các giá trị chuỗi thời gian.

3. Xây dựng mối quan hệ mờ và xác định nhóm các quan hệ logic mờ theo **Định nghĩa 7**.

4. Dự báo chuỗi thời gian mờ theo các luật sau:

*Luật 1:* Nếu nhóm quan hệ mờ  $A_i \rightarrow \emptyset$  thì giá trị dự báo mờ tại thời điểm  $t$  sẽ là  $A_i$

*Luật 2:* Nếu nhóm quan hệ logic mờ có dạng  $A_i \rightarrow A_k$  giá trị dự báo mờ tại thời điểm  $t$  sẽ là  $A_k$

*Luật 3:* Nếu nhóm mỗi quan hệ mờ phụ thuộc thời gian có dạng  $A_i \rightarrow A_{i1}, A_{i2} \dots A_{ip}$ , thì giá trị dự báo sẽ là:  $A_{i1}, A_{i2} \dots A_{ip}$

5. Giải mờ dựa vào các luật dự báo:

*Luật 1:* Nếu nhóm quan hệ mờ của là rỗng khi đó giá trị dự báo của  $F(t)$  là giá trị  $A_i$  và giải mờ sẽ là điểm giữa của khoảng  $u_i$

$$\text{forecast} = m_i$$

*Luật 2:* Nếu nhóm quan hệ logic mờ có dạng  $A_i \rightarrow A_k$  và nếu điểm giữa của khoảng  $u_k$  là  $m_k$  thì

$$\text{forecast} = m_k$$

*Luật 3:* Nếu mỗi quan hệ mờ bậc cao có dạng  $A_{i2} \rightarrow A_{i1}, A_{i2} \dots A_{ip}$ , thì giá trị dự báo sẽ là:

$$\text{forecast} = \frac{1 \times m_{i1} + 2 \times m_{i2} + \dots + k \times m_{ik}}{1 + 2 + \dots + k}$$

với  $m_{i1}, m_{i2}, \dots, m_{ip}$  là điểm giữa của các đoạn tương ứng.

### 3. Dự báo số lượng sinh viên nhập học

Để xem xét tính hiệu quả của định nghĩa mới về nhóm quan hệ logic mờ, chúng tôi sử dụng dữ liệu của bài báo Chen [4] về số lượng học sinh nhập học của Trường đại học Alabama. Chúng tôi cũng sử dụng Mô hình chuỗi thời gian mờ bậc nhất của Chen khi thực hiện tính toán với cách xây dựng nhóm quan hệ logic mờ truyền thống và so sánh với kết quả của cùng mô hình với cách xây dựng.

**Bảng 1.** Số lượng sinh viên nhập học

Năm	Số sinh viên	Năm	Số sinh viên
1971	13055	1982	15433
1972	13563	1983	15497

1973	13867	1984	15145
1974	14696	1985	15163
1975	15460	1986	15984
1976	15311	1987	16859
1977	15603	1988	18150
1978	15861	1989	18970
1979	16807	1990	19328
1980	16919	1991	19337
1981	16388	1992	18876

Thuật toán cải tiến cho chuỗi thời gian mờ bao gồm các bước sau đây:

**Bước 1.** Xây dựng tập nền U. Xác định giá trị lớn nhất và nhỏ nhất của chuỗi thời gian trên là 19337 và 13055 sinh viên. Do vậy tập nền U được xác định là giá trị trong khoảng [13000, 20000]. Ta sẽ chia U thành 7 khoảng  $u_1, u_2, \dots, u_7$  với độ rộng là 1000 như trong [4], như vậy các khoảng sẽ là:  $u_1 = [13000, 14000]$ ,  $u_2 = [14000, 15000]$ ,  $\dots$ ,  $u_7 = [19000, 20000]$ .

**Bước 2:** Xây dựng các tập mờ xác định trên các biến ngôn ngữ là các khoảng đã chia

Trong bước này ta xác định lại các tập mờ  $A_i$  tương ứng với từng khoảng và có thể gán lại các giá trị ngôn ngữ cho từng tập mờ này. Các tập mờ  $A_i$   $i=1,2,\dots,7$  được định nghĩa thông qua các hàm thuộc để đơn giản có dạng hình nón nhận 3 giá trị 0, 0.5 và 1 và được viết như sau:

$$A_1 = 1/u_1 + 0.5/u_2 + 0/u_3 + \dots + 0/u_6 + 0/u_7$$

$$A_2 = 0.5/u_1 + 1/u_2 + 0.5/u_3 + \dots + 0/u_6 + 0/u_7$$

.....

$$A_6 = 0/u_1 + 0./u_2 + \dots + 0.5/u_5 + 1/u_6 + 0.5/u_7$$

$$A_7 = 0/u_1 + 0/u_2 + \dots + 0/u_5 + 0.5/u_6 + 1/u_7$$

**Bước 3.** Xác định mối quan hệ mờ và nhóm quan hệ mờ phụ thuộc thời gian

Theo định nghĩa phần trên ta lập chuỗi thời gian mờ tương ứng với các tập mờ ở trên và xác định mối quan hệ mờ tại thời điểm  $t=1,2,\dots,7$ . Có thể thấy ngay được các mối quan hệ đầu tiên như sau:  $A_1 \rightarrow A_1, A_1 \rightarrow A_1, A_1 \rightarrow A_2, \dots, A_7 \rightarrow A_6$ .

Từ đây xác định nhóm các mối quan hệ mờ theo **Định nghĩa 7** ở phần trên. Thí dụ ta có thể nhận được một nhóm quan hệ mờ liên quan đến vế trái  $A_3$  nhưng tại thời điểm khác nhau  $t=7, t=8, t=9$  ta lại có nhóm quan hệ logic mờ khác nhau:  $A_3(7) \rightarrow A_3, A_3, A_3(8) \rightarrow A_3, A_3, A_3; A_3(9) \rightarrow A_3, A_3, A_3, A_4$ . Toàn thể các nhóm quan hệ mờ sẽ được thể hiện dưới Bảng 2.

**Bảng 2.** Các nhóm mối quan hệ mờ

Giá Trị	Thời điểm	Giá Trị mờ	Nhóm QH mờ Chen	Nhóm QHLG mờ Yu	Nhóm QH logic mờ mới
13055	t=1	A1			
13563	t=2	A1	A1,A2	A1,A1,A2	A1
13867	t=3	A1	A1,A2	A1,A1,A2	A1,A1
14696	t=4	A2	A1,A2	A1,A1,A2	A1,A1,A2
15460	t=5	A3	A3	A3	A3
15311	t=6	A3	A3,A4	A3,A3,A3,A4,A3,A3,A3,A3,A4	A3
15603	t=7	A3	A3,A4	A3,A3,A3,A4,A3,A3,A3,A3,A4	A3,A3
15861	t=8	A3	A3,A4	A3,A3,A3,A4,A3,A3,A3,A3,A4	A3,A3,A3
16807	t=9	A4	A3,A4	A3,A3,A3,A4,A3,A3,A3,A3,A4	A3,A3,A3,A4
16919	t=10	A4	A3,A4,A6	A4,A4,A3,A6	A4
16388	t=11	A4	A3,A4,A6	A4,A4,A3,A6	A4,A4
15433	t=12	A3	A3,A4,A6	A4,A4,A3,A6	A4,A4,A3
15497	t=13	A3	A3,A4	A3,A3,A3,A4,A3,A3,A3,A3,A4	A3,A3,A3,A4,A3
1514	t=14	A3	A3,A4	A3,A3,A3,A4,A3,A3,A3,A3,A4	A3,A3,A3,A4,A3,A3