

**ĐẠI HỌC THÁI NGUYÊN  
TRƯỜNG ĐẠI HỌC CNTT VÀ TRUYỀN THÔNG**

**PHẠM THỊ NGÂN**

**MÔ HÌNH CHUỖI THỜI GIAN MỜ  
BẬC CAO HAI NHÂN TỐ VÀ ỨNG DỤNG**

**LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH**

**THÁI NGUYÊN - 2012**

## MỤC LỤC

MỤC LỤC.....	1
DANH MỤC BẢNG BIỂU .....	5
DANH MỤC HÌNH VẼ.....	6
MỞ ĐẦU.....	7
CHƯƠNG 1: TỔNG QUAN VỀ LÝ THUYẾT TẬP MỜ .....	10
1.1 Lý thuyết tập mờ.....	10
1.1.1 Tập mờ .....	10
1.1.2 Các phép toán trên tập mờ .....	11
1.1.2.1 Phần bù của tập mờ .....	11
1.1.2.2 Phép giao hai tập mờ .....	11
1.1.2.3. Phép hợp hai tập mờ.....	13
1.1.2.4. Luật De Morgan .....	14
1.1.2.5. Phép kéo theo .....	14
1.2 Các quan hệ và suy luận xấp xỉ, suy diễn mờ.....	14
1.2.1 Quan hệ mờ.....	14
1.2.1.1 Khái niệm về quan hệ rõ.....	14
1.2.1.2 Các quan hệ mờ .....	15
1.2.2 Suy luận xấp xỉ và suy diễn mờ .....	16
1.3 Hệ mờ .....	17
1.3.1 Bộ mờ hoá.....	18
1.3.2 Hệ luật mờ.....	18
1.3.4 Bộ giải mờ.....	19
CHƯƠNG 2: MÔ HÌNH CHUỖI THỜI GIAN MỜ .....	22
VÀ CÁC THUẬT TOÁN CƠ BẢN .....	22
2.1 Các kiến thức cơ bản về chuỗi thời gian .....	22
2.1.1 Khái niệm chuỗi thời gian .....	22
2.1.2 Tính chất của chuỗi thời gian .....	22
2.1.2.1 Tính dừng.....	22
2.1.2.2 Tuyến tính .....	23

2.1.2.3 Tính xu hướng.....	24
2.1.2.4 Tính mùa vụ .....	24
2.1.3 Phân loại chuỗi thời gian .....	24
2.1.3.2 Chuỗi thời gian tuyến tính.....	25
2.1.3.2 Chuỗi thời gian phi tuyến.....	25
2.1.3.3 Chuỗi thời gian đơn biến.....	25
2.1.3.4 Chuỗi thời gian đa biến.....	25
2.1.3.5 Chuỗi thời gian hỗn loạn.....	26
2.1.4 Mô hình chuỗi thời gian.....	26
2.2. Chuỗi thời gian mờ.....	27
2.2.1. Khái niệm.....	27
2.2.2. Một số định nghĩa liên quan đến chuỗi thời gian mờ .....	27
2.3. Một số thuật toán trong mô hình chuỗi thời gian mờ.....	28
2.3.1. Một số thuật toán bậc một (thuật toán cơ sở) .....	28
2.3.1.1. Thuật toán của Song & Chissom [7] .....	28
2.3.1.2. Thuật toán của Chen [10] .....	29
2.3.1.3. Thuật toán Heuristic của Huarng [12].....	30
2.3.2. Thuật toán bậc cao .....	31
2.3.3. Thuật toán chuỗi thời gian mờ có trọng .....	32
2.3.3.1 Mô hình chuỗi thời gian mờ có trọng của Yu .....	32
2.3.3.2 Thuật toán cải biên mô hình chuỗi thời gian mờ có trọng [5].....	33
2.3.3.3 Áp dụng dự báo số lượng sinh viên nhập học .....	35
<b>CHƯƠNG 3: MÔ HÌNH CHUỖI THỜI GIAN MỜ BẬC CAO HAI NHÂN TỐ VÀ TÍNH TOÁN THỬ NGHIỆM.....</b>	<b>39</b>
3.1 Khái niệm chuỗi thời gian mờ bậc cao .....	39
3.1.1 Khái niệm chuỗi thời gian mờ bậc cao một nhân tố .....	39
3.1.2 Khái niệm chuỗi thời gian mờ bậc cao hai nhân tố .....	40
3.2 Thuật toán trong mô hình chuỗi thời gian mờ bậc cao hai nhân tố.....	40
3.3 Ứng dụng trong dự báo.....	43
3.3.1 Ứng dụng thuật toán hai nhân tố bậc 2 .....	43

3.3.1.1 Dự báo nhiệt độ .....	43
3.3.1.2 Dự báo chỉ số chứng khoán .....	50
3.3.2 Ứng dụng thuật toán hai nhân tố bậc 3 .....	61
3.3.2.1 Dự báo nhiệt độ .....	61
3.3.2.2 Dự báo chỉ số chứng khoán .....	64
KẾT LUẬN .....	71
TÀI LIỆU THAM KHẢO .....	72

## DANH MỤC BẢNG BIỂU

Bảng 1.1 Một số phép kéo theo mờ thông dụng .....	14
Bảng 2.1 Số lượng sinh viên nhập học.....	35
Bảng 2.2 Các nhóm mối quan hệ mờ.....	36
Bảng 2.3 Kết quả dự báo của các phương pháp khác nhau .....	37
Bảng 2.4 So sánh hiệu quả thuật toán .....	38
Bảng 3.1 Chuỗi dữ liệu nhiệt độ trung bình hàng ngày từ ngày 01/06/2012 đến ngày 30/06/2012 lúc 7h sáng tại Hà Nội [3,4] (đơn vị tính: °C) .....	43
Bảng 3.2 Chuỗi dữ liệu độ che phủ của mây từ ngày 01/06/2012 đến ngày 30/06/2012 lúc 7h sáng tại Hà Nội (đơn vị tính: %).....	44
Bảng 3.3 Mờ hóa các giá trị nhiệt độ và độ che phủ của mây .....	46
Bảng 3.4 Mối quan hệ mờ hai nhân tố bậc 2 .....	47
Bảng 3.5 Kết quả dự báo nhiệt độ trung bình ngày .....	48
Bảng 3.6 Giá trị chỉ số chứng khoán TAIFEX.....	50
Bảng 3.7 Giá trị chỉ số chứng khoán TAIEX.....	51
Bảng 3.8 Mờ hóa các giá trị chỉ số TAIFEX và giá trị mờ của chỉ số TAIEX.....	54
Bảng 3.9: Mối quan hệ mờ hai nhân tố bậc 2 .....	55
Bảng 3.10 Kết quả dự báo chỉ số chứng khoán TAIFEX .....	58
Bảng 3.11 Mối quan hệ mờ hai nhân tố bậc 3 .....	61
Bảng 3.12 Kết quả dự báo nhiệt độ trung bình ngày .....	62
Bảng 3.13 Mối quan hệ mờ hai nhân tố bậc 3 .....	64
Bảng 3.14 Kết quả dự báo chỉ số chứng khoán TAIFEX .....	67

## DANH MỤC HÌNH VẼ

Hình 1.1 Hàm liên thuộc của tập mờ “x gần 1” .....	11
Hình 1.2 Giao của hai tập mờ.....	12
Hình 1.3 Phép hợp của hai tập mờ .....	13
Hình 1.4 Cấu hình cơ bản của hệ mờ.....	18
Hình 3.1 Đồ thị so sánh kết quả dự báo nhiệt độ và giá trị thực (bậc 2) .....	50
Hình 3.2 Đồ thị so sánh kết quả dự báo chỉ số chứng khoán và giá trị thực.....	61
Hình 3.3 Đồ thị so sánh kết quả dự báo nhiệt độ và giá trị thực (bậc 3) .....	64
Hình 3.4 Đồ thị so sánh kết quả dự báo và giá trị thực theo các thuật toán.....	69

## MỞ ĐẦU

Chuỗi thời gian đang được sử dụng như một công cụ hữu hiệu để phân tích số liệu trong kinh tế, xã hội cũng như trong nghiên cứu khoa học. Chính do tầm quan trọng của phân tích chuỗi thời gian, rất nhiều tác giả đã đề xuất các công cụ phân tích chuỗi thời gian để trích xuất ra những thông tin quan trọng từ trong các dãy số liệu.

Trước đây, phương pháp chủ yếu để phân tích chuỗi thời gian là sử dụng các công cụ của thống kê như hồi qui, phân tích Fourie và một vài công cụ khác. Nhưng hiệu quả nhất có lẽ là phương pháp sử dụng mô hình ARIMA của Box-Jenkins. Mô hình này đã cho một kết quả khá tốt trong phân tích dữ liệu và đang được sử dụng rất rộng rãi trong thực tế. Tuy nhiên trong một số lĩnh vực nhất là trong kinh tế, mô hình ARIMA chưa thể hiện tính hiệu quả vì chuỗi số liệu diễn biến mang tính chất phi tuyến. Do đó để dự báo chuỗi thời gian trong kinh tế, người ta phải có những cải biên như sử dụng mô hình ARCH. Tuy vậy vẫn còn khá nhiều hạn chế khi áp dụng mô hình này khi chuỗi số liệu ngắn và có nhiều biến động mang tính chất phi tuyến.

Để vượt qua được những khó khăn trên, gần đây nhiều tác giả đã sử dụng mô hình chuỗi thời gian mờ. Khái niệm tập mờ được Zadeh đưa ra từ năm 1965 và ngày càng tìm được ứng dụng trong nhiều lĩnh vực khác nhau nhất là trong điều khiển và trí tuệ nhân tạo. Trong lĩnh vực phân tích chuỗi thời gian, Song và Chissom [7-9] đã đưa ra khái niệm chuỗi thời gian mờ không phụ thuộc vào thời gian (chuỗi thời gian dừng) và phụ thuộc vào thời gian (không dừng) để dự báo. Chen [10-11] đã cải tiến và đưa ra phương pháp mới đơn giản và hữu hiệu hơn so với phương pháp của Song và Chissom. Trong phương pháp của mình, thay vì sử dụng các phép tính tổ hợp Max-Min phức tạp, Chen đã tính toán bằng các phép tính số học đơn giản để thiết lập các mối quan hệ mờ. Phương pháp của Chen cho hiệu quả cao hơn về mặt sai số dự báo và giảm độ phức tạp của thuật toán.

Từ những công trình ban đầu về mô hình chuỗi thời gian mờ, hiện nay số lượng công trình trong lĩnh vực này tăng lên rất nhanh và hiện nay cũng vẫn đang

được tiếp tục nghiên cứu. Mô hình đang được sử dụng để dự báo trong rất nhiều lĩnh vực của kinh tế hay xã hội như dự báo số sinh viên nhập trường, số khách du lịch, dân số, chứng khoán và trong đời sống như dự báo mức tiêu thụ điện, hay dự báo nhiệt độ của thời tiết... Tuy nhiên xét về độ chính xác của dự báo, các thuật toán trên cho kết quả chưa cao. Chính vì vậy, các nghiên cứu tập trung vào các mô hình khác nhau để nâng cao hiệu quả và độ chính xác của thuật toán.

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 phân cụm như công trình của Chen et al trong [10], tập thô hay sử dụng khái niệm tối ưu đám đông như trong công trình [11] để xây dựng các thuật toán trong mô hình chuỗi thời gian mờ. Một số tác giả đã sử dụng cả mạng nơ ron và giải thuật tiến hoá để xây dựng các hàm thuộc cũng như mối quan hệ mờ trong mô hình.

Một trong các hướng được tập trung phát triển là sử dụng mối quan hệ mờ bậc cao trong mô hình chuỗi thời gian mờ. Chen [10] tiếp tục là người đi đầu khi xây dựng được thuật toán để xử lý mối quan hệ mờ bậc cao. Sau đó hướng này được một số tác giả khác tiếp cận và ứng dụng trong các công trình của mình (Xem [3], [12-13]). Trong các công trình này, các tác giả chủ yếu sử dụng thuật toán của Chen nhưng có cải tiến đôi chút trong việc đưa ra các luật khác nhau để giải mờ.

Trong nghiên cứu các dãy số liệu tạo thành chuỗi thời gian, người ta nhận thấy rằng số liệu trong chuỗi thời gian chính có ảnh hưởng bởi nhiều thông tin khác. Thí dụ chỉ số chứng khoán của Đài Loan, Hàn Quốc hay Việt Nam đều bị ảnh hưởng bởi chỉ số chứng khoán Mỹ. Dãy số liệu đo nhiệt độ của một thành phố bị ảnh hưởng lớn bởi mức độ che phủ của mây. Điều này làm nảy sinh ra ý tưởng khi dự báo dãy số liệu chính có xét thêm một hay nhiều dãy số liệu phụ. Từ đó nảy sinh ra mô hình chuỗi thời gian mờ loại 2 khi đồng thời với chuỗi thời gian chính còn sử dụng số liệu của các tham số phụ để đưa ra dự báo [14-15].



Như đã trình bày ở trên, mô hình chuỗi thời gian mờ đang có nhiều ứng dụng trong công tác dự báo. Tuy nhiên kết quả dự báo của các phương pháp đề xuất còn chưa cao. Do đó việc tìm tòi các mô hình có độ chính xác cao hơn và thuật toán đơn giản hơn đang là một ưu tiên. Trong những năm gần đây một số 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 công trình của Chen và Hsu, Huarng, Singh,... Mô hình chuỗi thời gian mờ bậc cao đã được xem xét nhiều và được coi là một công cụ đắc lực để nâng cao hiệu quả tính toán. Cách tiếp cận khác là sử dụng mô hình chuỗi thời gian mờ bậc cao hai nhân tố đã được một số tác giả nghiên cứu hứa hẹn thu được nhiều kết quả tốt.

Với mục tiêu tìm hiểu về việc sử dụng mô hình chuỗi thời gian mờ trong dự báo, đặc biệt là việc sử dụng mô hình chuỗi thời gian mờ bậc cao hai nhân tố, em đã lựa chọn đề tài “**Mô hình chuỗi thời gian mờ bậc cao hai nhân tố và ứng dụng**” làm đề tài cho luận văn tốt nghiệp của mình.

Mục tiêu nghiên cứu của đề tài là tìm hiểu và nghiên cứu những khái niệm, tính chất và thuật toán trong mô hình chuỗi thời gian và đặt trọng tâm vào tìm hiểu mô hình chuỗi thời gian mờ bậc cao hai nhân tố và thử nghiệm tính hiệu quả của mô hình trong dự báo chỉ số chứng khoán, nhiệt độ. Luận văn được chia làm 3 chương:

Chương 1: Tổng quan về lý thuyết tập mờ.

Chương 2: Mô hình chuỗi thời gian mờ và các thuật toán cơ bản.

Chương 3: Mô hình chuỗi thời gian mờ bậc cao hai nhân tố và tính toán thử nghiệm.

Luận văn này được hoàn thành dưới sự hướng dẫn tận tình của TS Nguyễn Công Điều, tác giả xin bày tỏ lòng biết ơn chân thành của mình đối với thầy. Tác giả xin chân thành cảm ơn các thầy, cô giáo Viện Công nghệ thông tin, Trường Đại học Công nghệ Thông tin và Truyền thông - Đại học Thái Nguyên đã tham gia giảng dạy, giúp đỡ em trong suốt quá trình học tập nâng cao trình độ kiến thức. Tuy nhiên vì điều kiện thời gian và khả năng có hạn nên luận văn không thể tránh khỏi những thiếu sót. Tác giả kính mong các thầy cô giáo và bạn đóng góp ý kiến để đề tài được hoàn thiện hơn.

## CHƯƠNG 1: TỔNG QUAN VỀ LÝ THUYẾT TẬP MỜ

Trong các bộ môn toán cơ bản, suy luận logic nguyên thủy hay logic rõ với hai giá trị đúng/sai hay 1/0 đã rất quen thuộc. Tuy nhiên, các suy luận này không đáp ứng được hầu hết các bài toán phức tạp nảy sinh trong thực tế như những bài toán trong lĩnh vực điều khiển tối ưu, nhận dạng hệ thống,... mà các dữ liệu không đầy đủ, không được định nghĩa một cách rõ ràng. Trong những năm cuối thập kỷ 20, một ngành khoa học mới đã được hình thành và phát triển mạnh mẽ đó là hệ mờ. Đây là hệ thống làm việc với môi trường không hoàn toàn xác định, với các tham số, các chỉ tiêu kinh tế kỹ thuật, các dự báo về môi trường sản xuất kinh doanh chưa hoặc khó xác định một cách thật rõ ràng, chặt chẽ. Khái niệm logic mờ được giáo sư Lofti A.Zadeh đưa ra lần đầu tiên vào năm 1965 tại Mỹ. Từ đó lý thuyết mờ đã được phát triển và ứng dụng rộng rãi.

Chương này tập trung trình bày một số kiến thức cơ bản về hệ mờ có liên quan tới mô hình chuỗi thời gian mờ sẽ được đề cập tới ở chương sau.

### 1.1 Lý thuyết tập mờ

#### 1.1.1 Tập mờ

Định nghĩa: Cho  $\Omega$  ( $\Omega \neq \emptyset$ ) là không gian nền, một tập mờ  $A$  trên  $\Omega$  được xác định bởi hàm thuộc (membership function):

$$\mu_A: \Omega \rightarrow [0,1]$$

$$0 \leq \mu_A(x) \leq 1$$

$\mu_A(x)$ : Chỉ độ thuộc (membership degree) của phần tử  $x$  vào tập mờ  $A$  (để cho đơn giản trong cách viết, sau này ta ký hiệu  $A(x)$  thay cho hàm  $\mu_A(x)$ )

Khoảng xác định của hàm  $\mu_A(x)$  là đoạn  $[0,1]$ , trong đó giá trị 0 chỉ mức độ không thuộc về còn giá trị 1 chỉ mức độ thuộc về hoàn toàn.

Như vậy tập mờ  $A$  hoàn toàn xác định trên tập các bộ đôi:

$$A = \{(x, \mu_A(x)) | x \in \Omega\}$$