

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ
TRUYỀN THÔNG

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

HỌC VIÊN: TRẦN XUÂN HÙNG

NGƯỜI HƯỚNG DẪN: TS. NGUYỄN CÔNG ĐIỀU

ĐỀ TÀI: CÁC PHƯƠNG PHÁP TÍNH TOÁN TRONG SỰ
BÁO CHUỖI THỜI GIAN MỞ

THÁI NGUYÊN, 2015

MỤC LỤC

MỞ ĐẦU	1
CHƯƠNG 1.....	5
MỘT SỐ KHÁI NIỆM VỀ LÝ THUYẾT TẬP MỜ.....	5
1.1 Lý thuyết tập mờ	5
1.1.1 Định nghĩa tập mờ.....	5
1.2.2 Suy luận xấp xỉ và suy diễn mờ	17
1.3.1 Bộ mờ hoá.....	22
1.3.2 Hệ luật mờ.....	22
1.3.4 Bộ giải mờ.....	24
CHƯƠNG 2.....	26
MÔ HÌNH CHUỖI THỜI GIAN MỜ VÀ CÁC THUẬT TOÁN CƠ BẢN.....	26
2.1 Các kiến thức cơ bản về chuỗi thời gian	26
2.1.2 Tính chất của chuỗi thời gian	26
2.1.2.1 Tính dừng	26
2.1.2.2 Tuyến tính.....	27
2.1.2.3 Tính xu hướng	28
2.1.2.4 Tính mùa vụ.....	28
2.1.3 Phân loại chuỗi thời gian	28
2.1.3.1 Chuỗi thời gian tuyến tính	29
2.1.3.2 Chuỗi thời gian phi tuyến	29
2.1.3.3 Chuỗi thời gian đơn biến	29
2.1.3.4 Chuỗi thời gian đa biến	30
2.1.3.5 Chuỗi thời gian hỗn loạn.....	30
2.1.4 Mô hình chuỗi thời gian	31

2.2	<i>Chuỗi thời gian mờ</i>	31
2.2.1	Khái niệm.....	31
2.2.2	Một số định nghĩa liên quan đến chuỗi thời gian mờ	32
2.3	Một số thuật toán trong mô hình chuỗi thời gian mờ	33
2.3.1	Các phương pháp chia khoảng.....	33
2.3.1.1	<i>Phương pháp lựa chọn ngẫu nhiên</i>	34
2.3.1.2	<i>Phương pháp độ dài dựa trên sự phân bố giá trị</i>	34
2.3.1.3	<i>Phương pháp độ dài dựa trên giá trị trung bình</i>	35
2.3.1.4	<i>Phương pháp dựa trên mật độ</i>	35
2.3.2	Mô hình thuật toán của Song và Chissom.....	35
2.3.3	Mô hình thuật toán của Chen.....	36
2.3.4	Mô hình chuỗi thời gian mờ đơn giản của Singh	37
2.3.5	Mô hình chuỗi thời gian mờ bậc cao của Singh	40
	CHƯƠNG 3.....	44
	ỨNG DỤNG TRONG TÍNH TOÁN THỬ NGHIỆM.....	44
3.1	Ứng dụng trong dự báo	44
3.1.1	Dự báo mức tiêu thụ điện bằng mô hình đơn giản của Singh	44
3.1.2	So sánh kết quả dự báo của phương pháp Singh đơn giản và bậc cao với các phương pháp khác	51
3.2	Đồ thị so sánh kết quả	53
3.2.1	Đồ thị so sánh của Chen và Singh đơn giản	53
3.2.2	Đồ thị so sánh Chen với Singh bậc cao.....	55
	KẾT LUẬN	56
	PHỤ LỤC.....	58
	Chương trình:	58
	Singh đơn giản.....	58

Sinh bậc cao.....	62
Tài liệu tham khảo.....	69

DANH MỤC HÌNH VẼ

<i>Hình 1.1 Hàm thuộc $\mu_A(x)$ có mức chuyển đổi tuyến tính.....</i>	6
<i>Hình 1.2 Hàm thuộc của tập B.</i>	7
<i>Hình 1.3 Miền xác định và miền tin cậy của tập mờ A.</i>	8
<i>Hình 1.4 Biểu diễn tập mờ chiều cao.</i>	9
<i>Hình 1.5 Tập bù \bar{A} của tập mờ A.</i>	10
<i>Hình 1.6 Hợp hai tập mờ có cùng tập vũ trụ.</i>	11
<i>Hình 1.7 Giao hai tập mờ có cùng tập vũ trụ.</i>	11
<i>Hình 1.8 Biểu diễn theo biểu đồ Sagittal.....</i>	16
<i>Hình 1.9 Cấu hình cơ bản của hệ mờ.....</i>	22

DANH MỤC BẢNG

<i>Bảng 1.1 Biểu diễn tập mờ A.....</i>	7
<i>Bảng 1.2 Một số phép kéo theo mờ thông dụng.....</i>	13
<i>Bảng 2.1 Ánh xạ cơ sở.....</i>	34
<i>Bảng 3.1 Số liệu mức độ tiêu thụ điện tại trường Cao đẳng Y tế Phú Thọ.....</i>	44
<i>Bảng 3.2 Phân bố giá trị trong từng khoảng.....</i>	46
<i>Bảng 3.3 Phân khoảng.....</i>	47
<i>Bảng 3.4 Mối quan hệ mờ.....</i>	48
<i>Bảng 3.5 Nhóm mối quan hệ mờ.....</i>	49
<i>Bảng 3.6 Kết quả dự báo của Chen.....</i>	50
<i>Bảng 3.7 Bảng so sánh kết quả dự báo.....</i>	51

DANH MỤC BIỂU ĐỒ

<i>Biểu đồ 3.1 Biểu đồ so sánh 1</i>	54
<i>Biểu đồ 3.2 Biểu đồ so sánh 2</i>	55

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 Fourier 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 [1-3] 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 [4] 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 các công trình ban đầu về chuỗi thời gian mờ được xuất hiện năm 1993, hiện nay mô hình này đang được sử dụng để dự báo trong rất nhiều lĩnh vực của kinh tế hay xã hội như giáo dục để dự báo số sinh viên nhập trường [2], [4] hay trong lĩnh vực dự báo thất nghiệp, 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. Để nâng cao độ chính xác của dự báo, một số thuật toán cho mô hình chuỗi thời gian mờ liên tiếp được đưa ra. Chen [5] đã sử dụng mô hình bậc cao của chuỗi thời gian mờ để tính toán. Sah và Degtiarev thay vì dự báo chuỗi thời gian đã sử dụng chuỗi thời gian là hiệu số bậc nhất để nâng cao độ chính xác và làm giảm độ phi tuyến.

Trong thời gian gần đây có khá nhiều cải tiến được các nhà nghiên cứu trên thế giới đưa ra để cải tiến độ chính xác của mô hình theo nhiều hướng khác nhau. Chen (2002) dựa trên mô hình trước đây đã đưa ra mô hình chuỗi thời gian mờ bậc cao và ứng dụng trong dự báo. Huarng (2001) đã nghiên cứu ảnh hưởng của độ dài khoảng lên độ chính xác của mô hình và đã đề xuất ra hai phương pháp chia khoảng là phân chia dựa trên phân bố và dựa trên giá trị trung bình. Tiếp theo hướng phát triển này, Huarng và Yu (2006), Chen và Chung (2006), Kuo (2008) đã tập trung vào việc phân chia khoảng để nâng cao độ chính xác của mô hình. Chen và Chung (2006) đã sử dụng giải thuật gen để điều chỉnh độ dài của khoảng cho mô hình bậc một và bậc cao của chuỗi thời gian mờ. Li và Cheng (2008) đã sử dụng thuật toán c-mean mờ cũng cho mục đích này. Cuối cùng là Kuo và các tác giả khác (2008) đã đề xuất thuật toán dựa trên phương pháp tối ưu đám đông để cải tiến cách xây dựng độ dài của khoảng.

Một hướng khác là sử dụng các cấu trúc khác nhau về mối quan hệ logic mờ để xây dựng các luật dự báo. Yu (2005) đã chú ý đến tính lặp lại của các tập mờ trong nhóm quan hệ logic mờ để gán tầm quan trọng của chúng bằng

các giá trị trọng số của mỗi lần lặp. Dieu N.C. (2010) đã chú ý đến yếu tố thời gian trong nhóm quan hệ logic mờ của Yu và đề xuất khái niệm nhóm quan hệ logic mờ phụ thuộc thời gian và ứng dụng trong dự báo.

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. Trong số các phương pháp cải tiến, mô hình của Singh đáng quan tâm chủ yếu đơn giản trong thuật toán nhưng cho hiệu quả cao trong thực tế. Đặc biệt các thuật toán đưa ra trong mô hình này rất thuận tiện cho việc lập trình.

Với mục tiêu tìm hiểu về việc sử dụng mô hình chuỗi thời gian mờ trong dự báo, em đã lựa chọn đề tài “**Các phương pháp tính toán trong dự báo chuỗi thời gian mờ**” mà trọng tâm là các mô hình tính toán của Singh. Các mô hình này đặt trọng tâm là xây dựng các công cụ tính toán khá đơn giản để dự báo và mô hình được xét cả mô hình chuỗi thời gian mờ bậc nhất và bậc cao. Sau đó em sử dụng các mô hình này để dự báo “**mức độ tiêu thụ điện tại trường cao đẳng Y tế Phú Thọ**” làm minh họa cho tính hiệu quả của các mô hình đã đề xuất trong luận văn tốt nghiệp của mình.

Với Mục tiêu trên, nội dung của đề tài là tìm hiểu và nghiên cứu những khái niệm, tính chất và thuật toán trong mô hình chuỗi thời gian mờ và đặt trọng tâm vào tìm hiểu **Các phương pháp tính toán trong dự báo chuỗi**

thời gian mờ của Singh và thử nghiệm tính hiệu quả của mô hình trong dự báo **mức độ tiêu thụ điện tại trường cao đẳng Y tế Phú Thọ**. Luận văn được chia làm 3 chương:

Chương 1: Một số khái niệm về lý thuyết tập mờ.

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

Chương 3: Ứng dụng trong tính toán thử nghiệm.

Luận văn này được hoàn thành dưới sự hướng dẫn tận tình của **TS Nguyễn Công Điều**, em xin bày tỏ lòng biết ơn chân thành của mình đối với thầy. Em 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. Em kính mong các thầy cô giáo và bạn đóng góp ý kiến để đề tài được hoàn thiện hơn.

CHƯƠNG 1

MỘT SỐ KHÁI NIỆM 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 Định nghĩa tập mờ

Tập mờ A xác định trên tập vũ trụ X là một tập mà mỗi phần tử của nó là một cặp các giá trị $(x, \mu_A(x))$, trong đó $x \in X$ và μ_A là ánh xạ:

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

Ánh xạ μ_A được gọi là hàm thuộc hoặc hàm liên thuộc (hoặc hàm thành viên - membership function) của tập mờ A . Tập X được gọi là cơ sở của tập mờ A .