

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

ĐẶNG THỊ THU THẢO

**MÔ HÌNH CHUỖI THỜI GIAN MỜ CÓ TRỌNG SỐ
BẬC CAO VÀ ỨNG DỤNG**

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

THÁI NGUYÊN - 2012

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

ĐẶNG THỊ THU THẢO

**MÔ HÌNH CHUỖI THỜI GIAN MỜ CÓ TRỌNG SỐ
BẬC CAO VÀ ỨNG DỤNG**

Chuyên ngành: Khoa học máy tính
Mã số: 60 48 01

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

NGƯỜI HƯỚNG DẪN KHOA HỌC: TS. NGUYỄN CÔNG ĐIỀU

THÁI NGUYÊN - 2012

LỜI CAM ĐOAN

Tôi cam đoan đây là công trình nghiên cứu của riêng tôi. Các số liệu, kết quả nêu trong luận văn là trung thực và chưa từng được ai công bố trong bất kỳ công trình nào khác.

Tác giả luận văn

Đặng Thị Thu Thảo

MỤC LỤC

Trang bìa phụ	
Lời cam đoan	
MỤC LỤC.....	3
DANH MỤC BẢNG BIỂU	iii
DANH MỤC HÌNH VẼ.....	iv
MỞ ĐẦU.....	1
CHƯƠNG 1 TỔNG QUAN VỀ CHUỖI THỜI GIAN	5
1.1 Các kiến thức cơ bản về chuỗi thời gian	5
1.1.1. <i>Khái niệm chuỗi thời gian</i>	5
1.1.2 <i>Tính chất của chuỗi thời gian</i>	6
1.1.3 <i>Phân chia chuỗi thời gian</i>	8
1.2 Mô hình chuỗi thời gian.....	10
1.3 Mô hình hồi quy	10
1.3.1 <i>Mô hình tự hồi quy (AR)</i>	11
1.3.2 <i>Mô hình trung bình trượt (MA)</i>	12
1.1.4 <i>Những hạn chế của mô hình ARMA trong chuỗi thời gian tài chính</i>	13
CHƯƠNG 2	14
MÔ HÌNH CHUỖI THỜI GIAN MỜ CÓ TRỌNG SỐ BẬC CAO	14
2.1 Tổng quan về tập mờ.....	14
2.1.1 <i>Tập mờ</i>	14
2.1.2 <i>Quan hệ mờ</i>	16
2.1.3 <i>Các phép toán của quan hệ mờ</i>	18
2.1.4 <i>Suy luận xấp xỉ và suy diễn mờ</i>	18
2.2 Hệ mờ.....	20
2.2.1 <i>Bộ mờ hoá</i>	20
2.2.2 <i>Giải mờ</i>	21
2.3 Chuỗi thời gian mờ.....	22

2.3.1 Một số khái niệm cơ bản	22
2.3.2 Một số định nghĩa liên quan đến chuỗi thời gian mờ	23
2.4 Một số thuật toán trong mô hình chuỗi thời gian mờ.....	24
2.4.1 Một số thuật toán bậc một (thuật toán cơ sở)	24
2.4.2 Một số thuật toán bậc cao.....	26
2.4.3 Chuỗi thời gian mờ có trọng bậc cao	30
2.4.4 Mô hình chuỗi thời gian mờ có trọng của Hui – Kuang Yu.....	32
2.4.5 Thuật toán bậc cao có trọng	38
CHƯƠNG 3	41
ỨNG DỤNG MÔ HÌNH CHUỖI THỜI GIAN MỜ	41
CÓ TRỌNG BẬC CAO.....	41
3.1 Ứng dụng trong bài toán dự báo nhiệt độ	41
3.1.1 Ứng dụng thuật toán chuỗi thời gian mờ có trọng bậc 3.....	41
3.1.2 Ứng dụng thuật toán chuỗi thời gian mờ có trọng bậc 1, bậc 2.....	45
3.2 Ứng dụng trong dự báo chỉ số Chứng khoán	48
3.2.1 Dự báo chỉ số chứng khoán Đài Loan	48
3.2.2 Dự báo chỉ số chứng khoán Việt Nam	56
KẾT LUẬN	62
TÀI LIỆU THAM KHẢO.....	64
PHỤ LỤC.....	66

DANH MỤC BẢNG BIỂU

Bảng 2.1 Số lượng sinh viên nhập học.....	34
Bảng 2.2 Các nhóm mối quan hệ mờ.....	35
Bảng 2.3 Kết quả dự báo của các phương pháp khác nhau	36
Bảng 2.4 So sánh hiệu quả thuật toán	37
Bảng 3.1 Nhiệt độ trung bình từ 01.6.1996 đến 30.9.1996.....	41
Bảng 3.2 Các giá trị mờ hóa.....	42
Bảng 3.3 Nhóm quan hệ mờ có trọng bậc 3.....	43
Bảng 3.4 Rút gọn của nhóm quan hệ mờ có trọng bậc 3	44
Bảng 3.5 Dự báo nhiệt độ trung bình của tháng 6.1996	44
Bảng 3.6 Dự báo nhiệt độ trung bình của tháng 6.1996 bằng mô hình bậc 1.....	45
Bảng 3.7 Dự báo nhiệt độ trung bình của tháng 6.1996 bằng mô hình bậc 2.....	46
Bảng 3.8 So sánh hiệu quả thuật toán	47
Bảng 3.9 Dữ liệu chỉ số chứng khoán TAIFEX.....	48
Bảng 3.10 Các giá trị mờ hóa.....	49
Bảng 3.11 Nhóm quan hệ mờ có trọng bậc 3.....	51
Bảng 3.12 Rút gọn của nhóm quan hệ mờ có trọng bậc 3	52
Bảng 3.13 Kết quả dự báo chỉ số chứng khoán TAIFEX	53
Bảng 3.14 So sánh với các phương pháp dự báo khác.....	54
Bảng 3.15 Số liệu chỉ số VN-index trong tháng 4 và tháng 5 năm 2012.....	56
Bảng 3.16 Các giá trị mờ hóa.....	57
Bảng 3.17 Nhóm quan hệ mờ có trọng bậc 3.....	59
Bảng 3.18 Rút gọn của nhóm quan hệ mờ có trọng bậc 3	60
Bảng 3.19 Giá trị dự báo	60

DANH MỤC HÌNH VẼ

Hình 2.1 Hàm liên thuộc của tập mờ “cao”	16
Hình 2.2 Cấu hình cơ bản của hệ mờ.....	20
Hình 2.3 Minh hoạ các phương pháp giải mờ.....	22
Hình 2.4 Đồ thị kết quả dự báo kết quả theo các thuật toán	38
Hình 3.1 Biểu đồ so sánh giá trị dự báo giữa các bậc	48
Hình 3.2 Biểu so sánh giá trị thực và giá trị dự báo.....	56
Hình 3.3 Biểu đồ so sánh giá trị thực tế và giá trị dự báo chỉ số VN-index	62
Hình PL.1. Giao diện chương trình.....	66
Hình PL.2. Chương trình dự báo nhiệt độ.....	67
Hình PL.3. Chương trình dự báo chỉ số VN-Index	67
Hình PL.4. Chương trình dự báo chỉ số chứng khoán	68

MỞ ĐẦU

Khoa học máy tính ngày nay phát triển luôn gắn liền với cuộc sống kinh tế xã hội. Nó không còn là việc lập trình ra những phần mềm quản lý để vận hành máy móc trong một số lĩnh vực cụ thể của ngành công nghệ thông tin đơn thuần. Giờ đây việc đi sâu vào tính ứng dụng với khả năng phân tích các số liệu trong kinh tế, xã hội một cách khoa học để có được những kết quả tính toán tối ưu đang trở thành một công cụ đắc lực giúp cho các nhà quản lý, các nhà đầu tư dự báo hay đánh giá được tính chính xác trong kết quả công việc của mình. Để có được những kết quả đánh giá tối ưu với tính chính xác cao từ kho dữ liệu tích lũy được, đòi hỏi các nhà khoa học phải luôn đi tìm các hướng tiếp cận để phân tích cũng như dự báo số liệu và phương pháp phân tích chuỗi thời gian đang là hướng đi mà các nhà khoa học lựa chọn và kỳ vọng.

Bằng các công cụ hữu hiệu của xác suất thống kê, phân tích chuỗi thời gian là công cụ quan trọng để phân tích số liệu trong kinh tế, xã hội cũng như trong nghiên cứu khoa học và từ đó trích xuất ra những thông tin quan trọng từ các dãy số liệu thống kê cơ bản.

Phương pháp phân tích chuỗi thời gian trước đây chủ yếu sử dụng các công cụ thống kê như hồi qui, phân Fourier và các công cụ phân tích khác nhưng kết quả đem lại chưa cao.

Phương pháp hiệu quả nhất có lẽ phải kể đến là phương pháp sử dụng mô hình ARIMA của Box-Jenkins. Ưu điểm của mô hình này là cho 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 lại chưa thể hiện được 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, các nhà khoa học phải có những cải biên như sử dụng mô hình ARCH để có được những phân tích và đánh giá về sự rủi ro gặp phải.

Để vượt qua được những khó khăn trên trong phân tích chuỗi thời gian, 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 [10-12] 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 [14] 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 thiết lập nhóm các mối quan hệ mờ và qua đó sử dụng các phép tính số học đơn giản để tính toán dự báo. 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 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.

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 [16], tập thô [4] hay sử dụng khái niệm tối ưu đám đông như trong công trình [8] để xây dựng các thuật toán trong mô hình chuỗi thời gian mờ. 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 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. Có thể kể ra đây công trình của Chu et.al [6].

Một trong các hướng được phát triển là sử dụng mối quan hệ mờ bậc cao trong mô hình chuỗi thời gian mờ. Chen [15] 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. 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. Riêng Singh trong bài báo [17] đã xây dựng mô hình chuỗi thời gian mờ bậc cao bằng cách mở rộng thuật toán đơn giản của mình xây dựng trong các công trình trước đây.

Một cải tiến trong mô hình bậc cao là dùng các trọng số để nâng cao độ chính xác thuật toán. Tư tưởng của nó là khi tạo các nhóm quan hệ mờ, nhiều khi không tính đến sự lặp lại của các tập mờ trong nhóm. Để tính đến sự đóng góp