

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

NGUYỄN NGỌC LINH

**XÁC ĐỊNH ĐỘ DÀI KHOẢNG TRONG MÔ HÌNH DỰ BÁO
CHUỖI THỜI GIAN MỜ**

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

Thái Nguyên – 2016

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

NGUYỄN NGỌC LINH

**XÁC ĐỊNH ĐỘ DÀI KHOẢNG TRONG MÔ HÌNH DỰ BÁO
CHUỖI THỜI GIAN MỜ**

Chuyên ngành : Khoa học máy tính

Mã số : 60 48 01 01

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

NGƯỜI HƯỚNG DẪN KHOA HỌC : TS. Nguyễn Công Điều

Thái Nguyên – 2016

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

Nguyễn Ngọc Linh

MỤC LỤC

	Trang
Trang phụ bìa	
MỤC LỤC.....	i
MỞ ĐẦU.....	1
CHƯƠNG 1: CÁC KIẾN THỨC CƠ BẢN VỀ TẬP MỜ.....	4
CHUỖI THỜI GIAN VÀ CHUỖI THỜI GIAN MỜ.....	4
1.1. Lý thuyết tập mờ [1]	4
1.1.1. Tập mờ	4
1.1.2. Các phép toán trên tập mờ	5
1.2. Các quan hệ và suy luận xấp xỉ, suy diễn mờ [1]	7
1.2.1. Quan hệ mờ	7
1.2.2. Suy luận xấp xỉ và suy diễn mờ	9
1.2.3. Bộ mờ hoá	10
1.2.4. Hệ luật mờ	11
1.2.5. Động cơ suy diễn	11
1.2.6. Bộ giải mờ	12
1.3. Chuỗi thời gian và quá trình ngẫu nhiên.....	13
1.3.1. Khái niệm chuỗi thời gian và quá trình ngẫu nhiên.....	13
1.3.2. Hàm tự tương quan	14
1.3.3. Quá trình ARMA	15
1.4. Các khái niệm về chuỗi thời gian mờ	19
1.4.1. Các định nghĩa	19
1.4.2 Một số mô hình chuỗi thời gian mờ.....	21
CHƯƠNG 2: CÁC PHƯƠNG PHÁP CHIA KHOẢNG CỦA HUANG....	25
2.1. Phương pháp tính độ dài phân bố	25
2.2. Phương pháp tính độ dài dựa trên mức trung bình	27
2.3. Phương pháp tính độ dài khoảng thời gian dựa theo tỷ lệ	28

CHƯƠNG 3: TÍNH TOÁN THỬ NGHIỆM VÀ CÁC ĐÁNH GIÁ	32
3.1. Tính toán thử nghiệm.....	32
3.2 Đánh giá các phương pháp chia khoảng	59
KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN CỦA ĐỀ TÀI	61
PHỤ LỤC.....	63

DANH MỤC HÌNH VẼ, BẢNG BIỂU, ĐỒ THỊ

Hình 1.1. Giao hai tập mờ có cùng không gian nền.....	6
Hình 1.2. Hợp hai tập mờ có cùng không gian nền	6
Bảng 1.1 Một số phép kéo theo mờ thông dụng	7
Hình 1.3. Hình vẽ biểu diễn theo biểu đồ sagital.....	8
Hình 1.4. Số lượng hàng hóa đã bán được trong siêu thị.....	13
Bảng 2.1 Cơ sở ánh xạ	26
Bảng 2.2 Bảng cơ sở	29
Bảng 3.1. Bảng số liệu chỉ số tiêu dùng cả nước trong tháng 1 từ	32
năm 1995 đến năm 2014	32
Bảng 3.2. Bảng tính hiệu số tuyệt đối theo phương pháp tính độ dài phân bố.....	33
Bảng 3.3. Bảng tính số lượng các hiệu số sai phân bậc 1 theo phương pháp tính độ dài phân bố	34
Bảng 3.4. Bảng mờ hóa và nhóm mối quan hệ mờ.....	35
Bảng 3.5. Giải mờ và kết quả dự báo theo phương pháp tính độ dài phân bố.....	36
Bảng 3.6. Bảng mờ hóa và nhóm mối quan hệ mờ.....	38
Bảng 3.7. Giải mờ và kết quả dự báo tính độ dài dựa trên mức trung bình. ..	38
Bảng 3.8. Kết quả tính toán sự khác biệt tương đối cho mỗi sự tuyệt đối.....	39
Bảng 3.9. Kết quả tính thứ tự r_i	39
Bảng 3.10. Mối quan hệ mờ và nhóm mối quan hệ mờ.....	41
Bảng 3.11. Kết quả dự báo theo phương pháp tính độ dài khoảng thời gian dựa theo tỷ lệ	41
Bảng 3.12. So sánh các phương pháp chia khoảng.....	42
Đồ thị 3.1 So sánh các kết quả dự báo chỉ số giá tiêu dùng của cả nước trong tháng 1 từ năm 1995 đến năm 2014	43
Bảng 3.13. Bảng số liệu chỉ số Vn-index trong tháng 9 và tháng 10 năm 2015.....	44

Bảng 3.14. Bảng tính hiệu số tuyệt đối theo phương pháp tính độ dài phân bố.....	45
Bảng 3.15. Bảng tính số lượng các hiệu sai số phân bậc 1 theo phương pháp tính độ dài phân bố.....	46
Bảng 3.16. Bảng mờ hóa và nhóm các mối quan hệ mờ	47
Bảng 3.17. Kết quả dự báo theo phương pháp tính độ dài phân bố.....	48
Bảng 3.18. Mờ hóa và nhóm mối quan hệ mờ.....	50
Bảng 3.19. Kết quả dự báo theo phương pháp tính độ dài trên mức trung bình.....	51
Bảng 3.20. Kết quả tính toán sự khác biệt tương đối cho mỗi sự tuyệt đối....	52
Bảng 3.21. Bảng kết quả tính thứ tự r_i	53
Bảng 3.22. Bảng thiết lập mối quan hệ mờ và nhóm mối quan hệ mờ.....	55
Bảng 3.23. Kết quả dự báo phương pháp tính độ dài khoảng thời gian	55
dựa theo tỷ lệ	55
Bảng 3.24. Bảng so sánh kết quả dự báo các phương pháp chia khoảng	56
Hình 3.2. Đồ thị so sánh các kết quả dự báo chỉ số VN-index với giá trị thực	58

MỞ ĐẦU

Chuỗi thời gian là một công cụ xử lý dữ liệu hữu hiệu trong thống kê. Mô hình được sử dụng rộng rãi nhất để phân tích chuỗi thời gian trong thống kê là mô hình ARMA do Box-Jenkins đề xuất. Tuy nhiên trên thực tế có khá nhiều chuỗi số liệu không thoả mãn được các tính chất để có thể xử lý được bằng công cụ thống kê như mô hình ARMA. Do vậy cần có những công cụ để có thể xử lý được những trường hợp đặc trưng này. Mô hình chuỗi thời gian mờ là một công cụ được phát triển nhằm đáp ứng nhu cầu này.

Chuỗi thời gian mờ và mô hình chuỗi thời gian mờ bậc nhất do Song và Chissom [9] phát triển từ năm 1993. Sau công trình này, nhiều 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...

Sự phát triển của mô hình chuỗi thời gian mờ có tiến bộ vượt bậc khi năm 1996 Chen [10] 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 mối quan hệ mờ. Phương pháp của Chen đã làm giảm khá nhiều về độ phức tạp của thuật toán. Nhiều công trình tiếp theo đã sử dụng cách tiếp cận này để dự báo cho chuỗi thời gian. Huarng 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.

Để nâng cao hiệu quả và độ chính xác của thuật toán, những năm gần đây đã có hàng loạt công trình đưa ra nhiều kỹ thuật khác nhau. Những công cụ trong lý thuyết tính toán mềm, khai phá dữ liệu, mạng nơ ron và các giải thuật tiến hoá, tối ưu bầy đàn đều được đưa vào sử dụng. Một trong các hướng được phát triển là sử dụng mối quan hệ mờ bậc cao trong mô hình chuỗi thời

gian mờ. Chen [11] tiếp tục là người đi đầu khi xây dựng được thuật toán để xử lý mối quan hệ mờ bậc cao. Các công trình tiếp theo của Chen đã đề cập đến khái niệm mô hình chuỗi thời gian mờ hai nhân tố kết hợp với mối quan hệ logic mờ bậc cao. Sau đó hướng này được một số tác giả khác tiếp cận và ứng dụng trong các công trình của mình.

Năm 2001 Huarng đã chú ý đến vai trò của việc phân khoảng trong mô hình chuỗi thời gian mờ. Ông nhận xét rằng: nhiều nhà khoa học đã cho thấy cách phân chia khoảng có ảnh hưởng rất lớn đến độ chính xác của thuật toán. Nếu phân các khoảng có độ dài lớn thì số phép tính giảm nhưng sẽ có sự phân tán kết quả, còn nếu chia khoảng nhỏ mất ý nghĩa của dự báo. Các tác giả có đề xuất nhiều cách khác nhau để phân khoảng như chia ngẫu nhiên, dựa vào giá trị trung bình, dựa vào phân bố hay dựa vào mật độ phân bố. Mỗi phương pháp được sử dụng trong các trường hợp khác nhau và đều cho kết quả tốt hơn so với phương pháp truyền thống. Như vậy cũng có thể thấy rõ sự ảnh hưởng của phương pháp chia khoảng đến kết quả dự báo. Ngoài ra còn có một số công trình khác đề xuất thêm các thuật toán chia khoảng khác nhau như trong các bài báo [2], [5],[6].

Dự báo chuỗi thời gian sử dụng mô hình chuỗi thời gian mờ có một số bước cơ bản như sau: Xác định tập nền, phân chia tập nền thành các khoảng, Mờ hoá các giá trị lịch sử, Xác định các mối quan hệ mờ, Dự báo và cuối cùng là giải mờ. Vấn đề tính độ dài của khoảng trong phân chia tập nền của thuật toán có ảnh hưởng lớn đến độ chính xác của mô hình dự báo. Đã xuất hiện khá nhiều phương pháp tính độ dài của khoảng của các tác giả khác nhau. Mặc dù vậy các thuật toán này có những ảnh hưởng khác nhau đến kết quả dự báo. Nghiên cứu ảnh hưởng các cách chọn khoảng của các thuật toán khác nhau đến độ chính xác của dự báo chuỗi thời gian cũng là một bài toán khá lý thú và cần có những đánh giá và tổng kết các phép phân chia này để sử dụng trong nhiều bài toán khác nhau. Chính vì lý do này, em đã lựa chọn đề tài “**Xác định**

độ dài khoảng trong mô hình dự báo chuỗi thời gian mờ” làm đề tài cho luận văn tốt nghiệp của mình.

Nội dung chính của luận văn có cấu trúc như sau:

Chương 1: Các kiến thức cơ bản về tập mờ, chuỗi thời gian và chuỗi thời gian mờ.

Chương 2: Các phương pháp chia khoảng của Huarng

Chương 3: Tính toán thử nghiệm và các đánh giá

Luận văn của em 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 cũng xin chân thành cảm ơn các thầy cô giáo Viện Công nghệ thông tin, trường Đại học Công nghệ thông tin và Truyền thông đã tham gia giảng dạy, giúp đỡ em trong suốt quá trình học tập nâng cao trình độ kiến thức. Mặc dù đã có nhiều cố gắng hoàn thiện luận văn bằng tất cả nhiệt huyết và năng lực của mình, tuy nhiên luận văn không thể tránh khỏi những thiếu sót, em rất mong nhận được sự đóng góp của quý thầy cô và các bạn./.