**GIẢI PHÁP PHÂN BIỆT TÊN NGƯỜI TRÊN WEB DỰA TRÊN MÔ HÌNH THÔNG TIN NGƯỜI VÀ THỬ NGHIỆM VÀO HỆ THỐNG TÌM KIẾM NGƯỜI TRONG TIẾNG VIỆT**

**Nguyễn Thị Thanh Na**

*Khóa QH-2007-I/CQ , ngành Công nghệ thông tin*

**Tóm tắt khóa luận tốt nghiệp:**

Tìm kiếm thông tin về thực thể người trên Web là một trong những hoạt động phổ biến nhất của người dùng trên Internet. Tuy nhiên, thực thể người là một trong những kiểu thực thể có độ nhập nhằng cao. Một tên có thể ứng với nhiều thực thể người khác nhau và nhiều tên có thể ứng với cùng một thực thể người. Bài toán phân biệt tên người nhằm giải quyết vấn đề nhập nhằng trong tên của thực thể người, từ đó nâng cao chất lượng tìm kiếm thông tin người trên Web.

Khóa luận này tập trung giải quyết bài toán phân biệt tên người trên Web dựa trên việc trích chọn đặc trưng theo mô hình thông tin Người. Các đặc trưng được trích chọn sử dụng để phân cụm các văn bản chứa cùng một tên người nhưng chỉ đến các thực thể người khác nhau.

Kết quả thực nghiệm ban đầu với tập tên người có độ nhập nhằng cao cho thấy mô hình phân biệt tên người của hệ thống đạt kết quả độ đo F0.5 = 84,8%% và F0,2 = 83.1 %. Điều này khẳng định mô hình là khả quan và có khả năng ứng dụng vào thực tế.

***Từ khóa:*** mô hình thông tin Người, phân biệt tên người, phân biệt nhập nhằng tên người.