

Một phương pháp lai trích xuất sự kiện và áp dụng vào hệ thống theo dõi tin tức trực tuyến $\mathcal{N}_{ewSOMoni}$

CÁN BỘ HƯỚNG DẪN

TS Phan Xuân Hiếu ThS. Trần Mai Vũ

SINH VIÊN THỰC HIỆN

Nguyễn Sỹ Quân Nguyễn Minh Hoàng
Ngô Quang Hiếu

Phòng thí nghiệm Công nghệ Tri thức – Đại học Công Nghệ



Nội dung

- 1 Giới thiệu
- 2 Phương pháp giải quyết
 - Các hướng tiếp cận
 - Phương pháp đề xuất
- 3 Hệ thống $\mathcal{N}_{ewSOMoni}$
 - Mô tả hệ thống
 - Demo hệ thống



Tài liệu tham khảo



Nội dung

- 1 Giới thiệu
- 2 Phương pháp giải quyết
 - Các hướng tiếp cận
 - Phương pháp đề xuất
- 3 Hệ thống $\mathcal{N}_{ewSOMoni}$
 - Mô tả hệ thống
 - Demo hệ thống



Động lực nghiên cứu

- Thông tin về một sự kiện nổi bật:
 - Nhiều
 - Đa chiều
 - Tính cập nhật cao
- Trích xuất sự kiện có vai trò quan trọng:
 - Hệ thống theo dõi, giám sát
 - BioCaster (<http://born.nii.ac.jp/>)
 - EpiSpider (<http://www.epispider.org/>)
 - Metro Monitor (www.metromonitor.com/)
 - Hệ thống cảnh báo, bảo mật
 - Frontex (<http://www.frontex.europa.eu/>)



Động lực nghiên cứu

- Thông tin về một sự kiện nổi bật:
 - Nhiều
 - Đa chiều
 - Tính cập nhật cao
- Trích xuất sự kiện có vai trò quan trọng:
 - Hệ thống theo dõi, giám sát
 - BioCaster (<http://born.nii.ac.jp/>)
 - EpiSpider (<http://www.epispider.org/>)
 - Metro Monitor (<http://www.metromonitor.com/>)
 - Hệ thống cảnh báo, bảo mật
 - Frontex (<http://www.frontex.europa.eu/>)



Động lực nghiên cứu

- Thông tin về một sự kiện nổi bật:
 - Nhiều
 - Đa chiều
 - Tính cập nhật cao
- Trích xuất sự kiện có vai trò quan trọng:
 - Hệ thống theo dõi, giám sát
 - BioCaster (<http://born.nii.ac.jp/>)
 - EpiSpider (<http://www.epispider.org/>)
 - Metro Monitor (<http://www.metromonitor.com/>)
 - Hệ thống cảnh báo, bảo mật
 - Frontex (<http://www.frontex.europa.eu/>)



Mô tả bài toán

- Trích xuất sự kiện theo phương pháp kết hợp luật ngữ nghĩa và học máy Maximum Entropy
- Áp dụng thử nghiệm cho hệ thống theo dõi tin tức
 - HÌNH SỰ
 - TAI NẠN GIAO THÔNG
 - CHÁY NỔ
- **Đầu vào:** một tin tức thu thập trực tuyến
- **Đầu ra:**
 - Bản tin chứa sự kiện không?
 - Thông tin sự kiện:
 - dạng sự kiện, tên sự kiện
 - nhân tố tham gia, thời gian, địa điểm



Mô tả bài toán

- Trích xuất sự kiện theo phương pháp kết hợp luật ngữ nghĩa và học máy Maximum Entropy
- Áp dụng thử nghiệm cho hệ thống theo dõi tin tức
 - HÌNH SỰ
 - TAI NẠN GIAO THÔNG
 - CHÁY NỔ
- **Đầu vào:** một tin tức thu thập trực tuyến
- **Đầu ra:**
 - Bản tin chứa sự kiện không?
 - Thông tin sự kiện:
 - dạng sự kiện, tên sự kiện
 - nhân tố tham gia, thời gian, địa điểm



Mô tả bài toán

- Trích xuất sự kiện theo phương pháp kết hợp luật ngữ nghĩa và học máy Maximum Entropy
- Áp dụng thử nghiệm cho hệ thống theo dõi tin tức
 - HÌNH SỰ
 - TAI NẠN GIAO THÔNG
 - CHÁY NỔ
- **Đầu vào:** một tin tức thu thập trực tuyến
- **Đầu ra:**
 - Bản tin chứa sự kiện không?
 - Thông tin sự kiện:
 - dạng sự kiện, tên sự kiện
 - nhân tố tham gia, thời gian, địa điểm



Mô tả bài toán

- Trích xuất sự kiện theo phương pháp kết hợp luật ngữ nghĩa và học máy Maximum Entropy
- Áp dụng thử nghiệm cho hệ thống theo dõi tin tức
 - HÌNH SỰ
 - TAI NẠN GIAO THÔNG
 - CHÁY NỔ
- **Đầu vào:** một tin tức thu thập trực tuyến
- **Đầu ra:**
 - Bản tin chứa sự kiện không?
 - Thông tin sự kiện:
 - dạng sự kiện, tên sự kiện
 - nhân tố tham gia, thời gian, địa điểm



Mô tả bài toán

- Trích xuất sự kiện theo phương pháp kết hợp luật ngữ nghĩa và học máy Maximum Entropy
- Áp dụng thử nghiệm cho hệ thống theo dõi tin tức
 - HÌNH SỰ
 - TAI NẠN GIAO THÔNG
 - CHÁY NỔ
- **Đầu vào:** một tin tức thu thập trực tuyến
- **Đầu ra:**
 - Bản tin chứa sự kiện không?
 - Thông tin sự kiện:
 - dạng sự kiện, tên sự kiện
 - nhân tố tham gia, thời gian, địa điểm



Mục tiêu nghiên cứu

- Thế nào là sự kiện, trích xuất sự kiện?
- Các hướng tiếp cận giải quyết bài toán trích xuất sự kiện?
- Làm thế nào để trích xuất sự kiện trên tin tức tiếng Việt?
- Một hệ thống theo dõi tin tức có khả thi không?



Nội dung

- 1 Giới thiệu
- 2 Phương pháp giải quyết
 - Các hướng tiếp cận
 - Phương pháp đề xuất
- 3 Hệ thống $\mathcal{N}_{ewSOMoni}$
 - Mô tả hệ thống
 - Demo hệ thống



Nội dung

- 1 Giới thiệu
- 2 Phương pháp giải quyết
 - Các hướng tiếp cận
 - Phương pháp đề xuất
- 3 Hệ thống $\mathcal{N}_{ewSOMoni}$
 - Mô tả hệ thống
 - Demo hệ thống



Định nghĩa sự kiện – trích xuất sự kiện



Áp dụng luật



Áp dụng học máy – thống kê



Kết hợp luật và học máy – thống kê



Nội dung

- 1 Giới thiệu
- 2 Phương pháp giải quyết
 - Các hướng tiếp cận
 - Phương pháp đề xuất
- 3 Hệ thống $\mathcal{N}_{ewSOMoni}$
 - Mô tả hệ thống
 - Demo hệ thống



Nội dung

- 1 Giới thiệu
- 2 Phương pháp giải quyết
 - Các hướng tiếp cận
 - Phương pháp đề xuất
- 3 Hệ thống $\mathcal{N}_{ewSOMoni}$
 - Mô tả hệ thống
 - Demo hệ thống

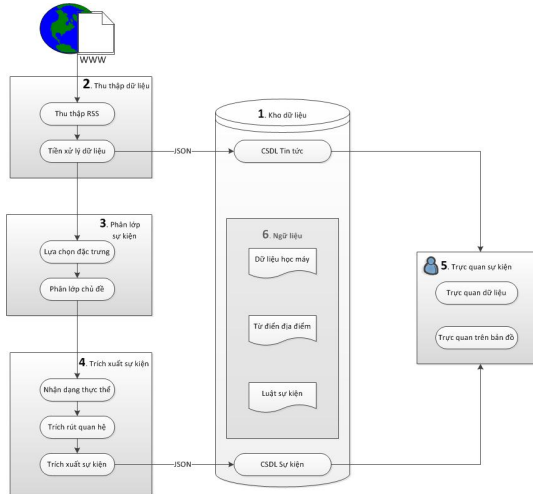


Nội dung

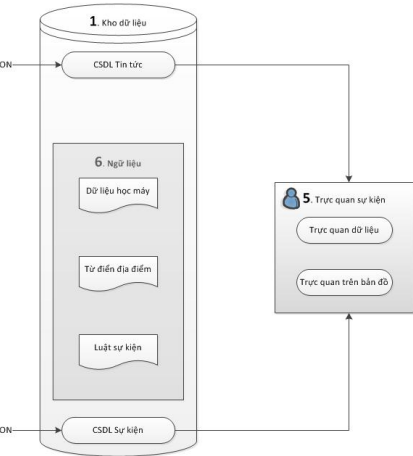
- 1 Giới thiệu
- 2 Phương pháp giải quyết
 - Các hướng tiếp cận
 - Phương pháp đề xuất
- 3 Hệ thống $\mathcal{N}_{ewSOMoni}$
 - Mô tả hệ thống
 - Demo hệ thống



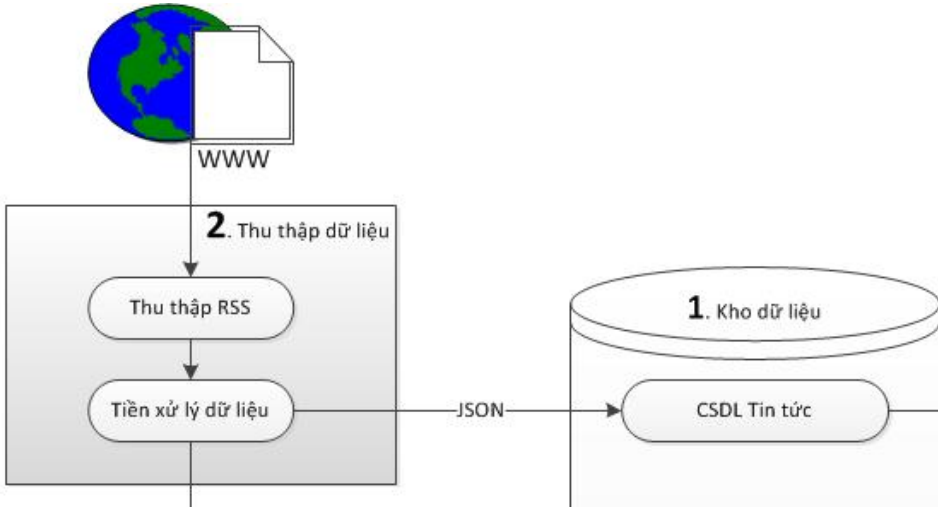
Mô hình hệ thống



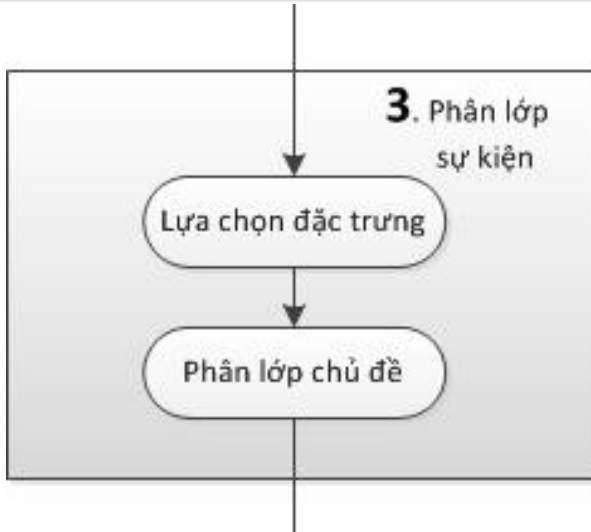
Mô hình hệ thống



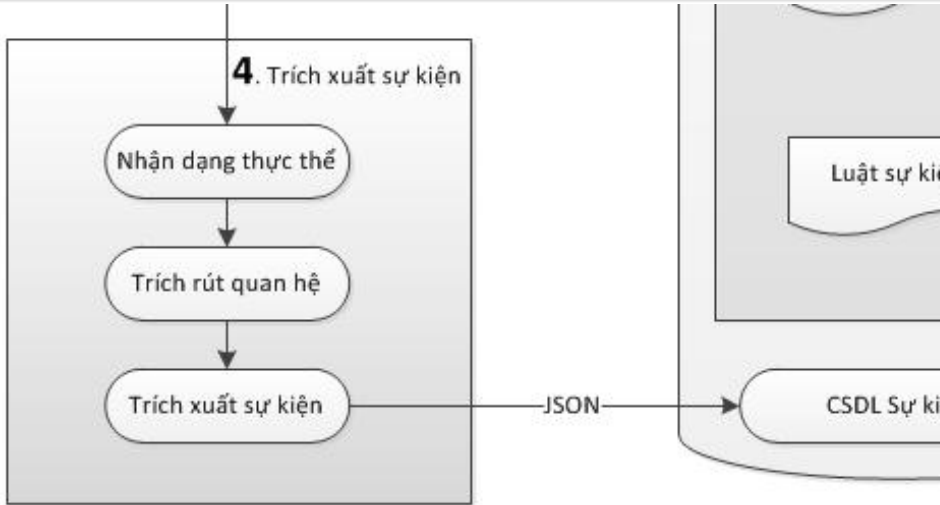
Mô hình hệ thống



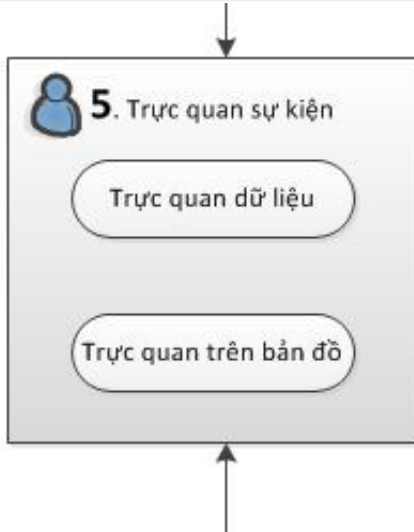
Mô hình hệ thống



Mô hình hệ thống



Mô hình hệ thống



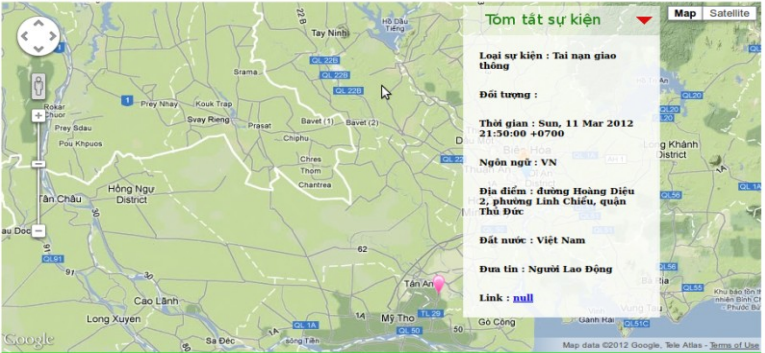
Nội dung

- 1 Giới thiệu
- 2 Phương pháp giải quyết
 - Các hướng tiếp cận
 - Phương pháp đề xuất
- 3 Hệ thống $\mathcal{N}_{ewSOMoni}$
 - Mô tả hệ thống
 - Demo hệ thống



Event Monitoring

Trang chủ | Giới thiệu | Xem: **Thành phố Hồ Chí Minh**



Tôm tắt sự kiện

Loại sự kiện : Tai nạn giao thông

Đối tượng :

Thời gian : Sun, 11 Mar 2012 21:50:00 +0700

Ngôn ngữ : VN

Địa điểm : đường Hoàng Diệu 2, phường Linh Chiểu, quận Thủ Đức

Đất nước : Việt Nam

Đưa tin : Người Lao Động

Link : null

Hỏa hoạn | Tai nạn giao thông | Hình sự

Tổng kết

- Các khái niệm cơ bản về mạng xã hội, phương tiện xã hội
- Đặc tính cơ bản của mạng xã hội
- Giám sát luồng thông tin trong mạng xã hội
- Ví dụ giám sát luồng thông tin qua phương tiện xã hội Twitter



CÂU HỎI???



CẢM ƠN THẦY VÀ CÁC BẠN
ĐÃ QUAN TÂM THEO DÕI!

