

Using Data Mining Techniques to Address Critical Information Exchange Needs in Disaster Affected Public-Private Networks

Li Zheng, Chao Shen, Liang Tang, Tao Li, Steve Luis,
Shu-Ching Chen, Vagelis Hristidis

School of Computer Science
Florida International University

Motivation: Hurricane Scenario

- Studies show that ~40% of companies *failed within 36 months* when they were closed for 3 or more days as a result of a hurricane.
- Communication network
 - Information exchanging
- Provide high-quality recovery info
 - Dynamic, summarized
- Facilitate collaboration
 - Personalized, community-based

The BCiN System

- A platform of information sharing, integration, extraction, and processing
- A data mining solution for disaster management and recovery
 - process and analyze the data from diverse and heterogeneous information sources with different formats
- An integration of advanced techniques
 - Natural Language Processing, Database, Data Mining

System Overview

Dashboard

Situation Dashboard Select Categories | See More Categories...

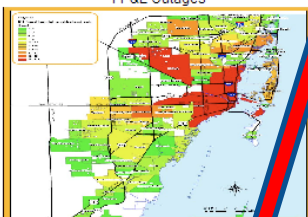
All Counties	Airport	County Buses	County Trains	Curfews	Drawbridges	PODS	Power Availability	Roadways	Schools	Shelters	Water ways	Seaport
Miami-Dade	⬅➡	●➡●	●	⬅➡		●●	⬅➡		⬅➡	●	●	⬅➡
Broward	●					●●	◆			●		⬅➡
Palm Beach	●	●		●	●	●	◆	⬅➡	●	●		
Monroe							◆	●				



Track/Timeline [Situation](#) [Messaging](#) [Reporting](#) [Resources](#) [Assistance](#) [ESF-18](#) [Special Request](#) [Directory](#) [Help](#)

Tracking

Impact: Miami-Dade



Event Timeline

All Counties

- Fri Nov 16, 2007 5:04 PM **Palm Beach** Martin Luther King Blvd. . (Fri Nov 16, 2007 3:04 PM)
Martin Luther king Blvd. is impassible. Find alternate routes
By: Butch Truesdale Palm Beach County Emergency Operations Center
- Fri Nov 16, 2007 5:05 PM **Broward** Port Everglades Fully Reopens (Fri Nov 16, 2007 1:05 PM)
Port Evergladesi has announced that it will return to full service November 17th. Passengers for cruise ships previously scheduled to leave before the 17th should contact their cruise line directly.
For more information please visit www.miamidade.gov
By: Frank Reddish Miami-Dade Department of Emergency Management and Homeland Security
- Fri Nov 16, 2007 8:00 PM **Miami-Dade** Curfews status (Thu Nov 15, 2007 4:00 PM)

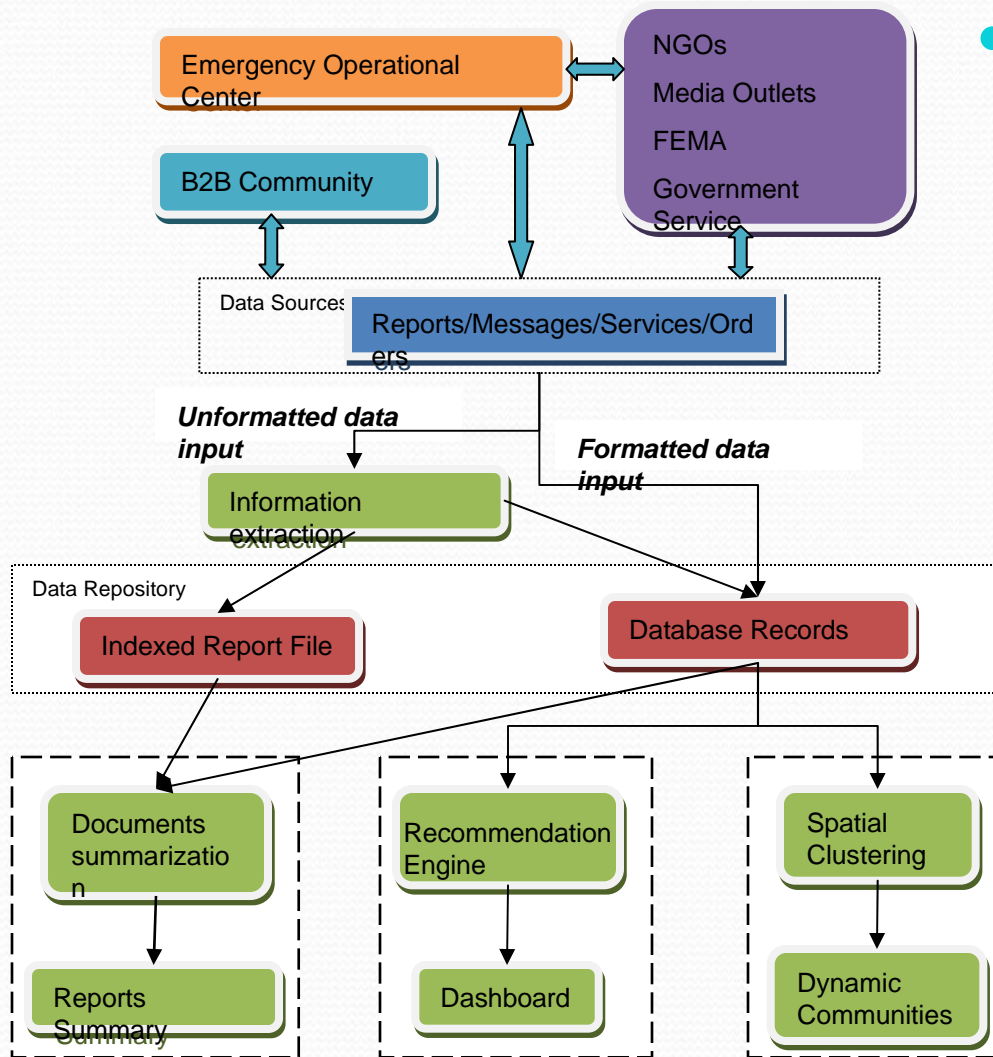
Reports submit/
File upload

Summarized reports

Challenges

- *How can the system quickly capture the status report information?*
 - *Information Extraction*
- *How can the system effectively understand the situation from a large collection of reports?*
 - *Multi-document Summarization*
- *How can we automatically capture users' interests and effectively deliver the relevant information to the users?*
 - *Dynamic dashboard*
- *How can we take advantage of the community information for disaster recovery?*
 - *Community generation*

Framework

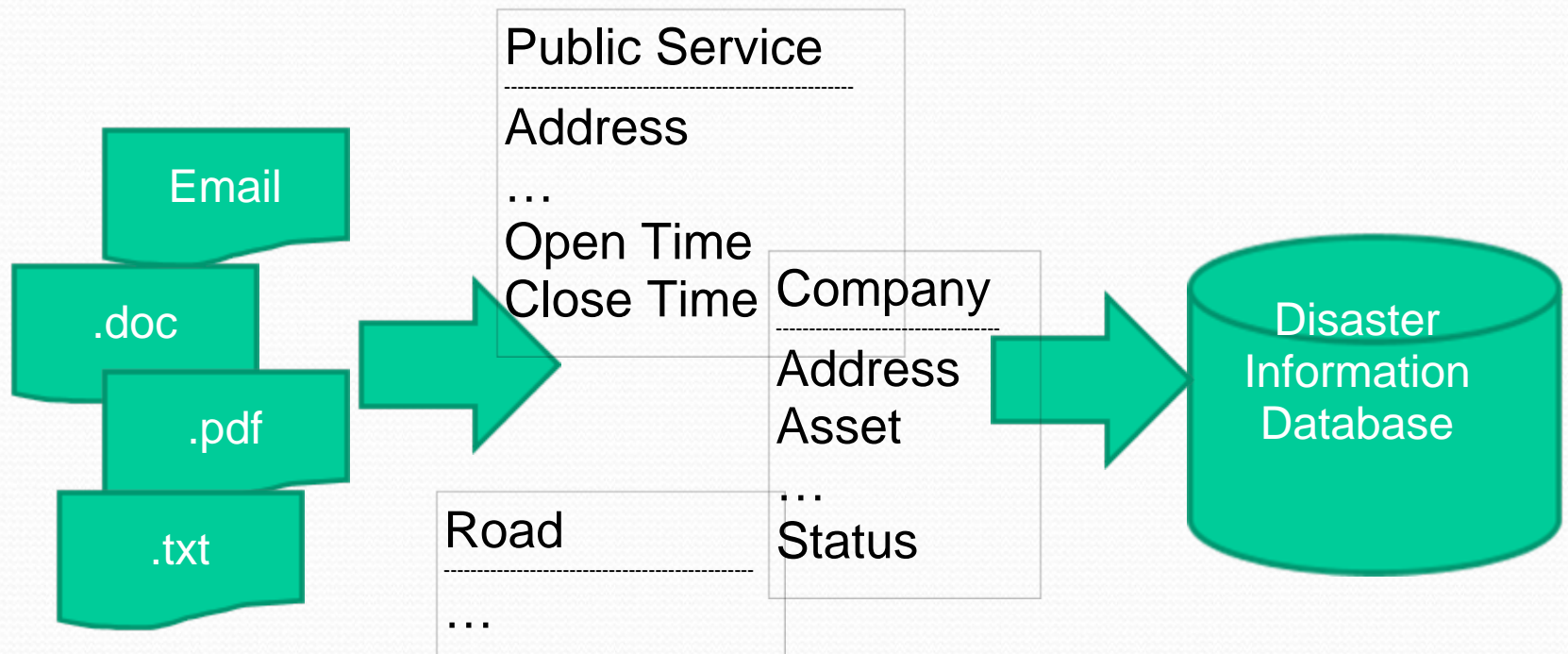


• Solutions

- Information extraction
- Content recommendation
- Reports summarization
- Community generation

Information Extraction

- Massive sources at granularities, redundancy
- Different varieties: structured and unstructured



Information extraction

Time: October 21, 2005 12:30 p.m.

Miami-Dade Emergency Operations Center is currently activated at a level II and officials and emergency managers are carefully monitoring Hurricane Wilma.

Residents are urged to finalize their personal hurricane preparations.

On Monday, October 24, Miami-Dade County offices, public schools, and courts will be closed.

Currently, transit bus and rail service continues, including Metrobus, Metrorail and Metromover.

Miami International Airport is open. However, if you have travel plans please check with your airline for flight information.

Tomorrow afternoon, the American Red Cross will open hurricane evacuation centers for residents who do not feel safe in their homes or live in low-lying areas.

- Information extraction
 - What was/is/will be the status of Facilities/Services/... at the time of...?
 - Unstructured and multi-source inputs.
 - Triple <Entity, Time, Status>.

Information extraction

Miami-Dade Emergency Operations Center is currently activated at a level II and officials and emergency managers are carefully monitoring Hurricane Wilma.

Residents are urged to finalize their personal hurricane preparations.

On **<T>Monday, October 24</T>**, **<E>Miami-Dade County offices</E>**, **<E>public schools</E>**, and **<E>courts</E>** will be closed.

<T>Currently</T>, **<E>transit bus</E>** and **<E>rail service</E>** continues, including **<E>Metrobus</E>**, **<E>Metrorail</E>** and **<E>Metromover</E>**.

<E>Miami International Airport</E> is open. However, if you have travel plans please check with your airline for flight information.

<T>Tomorrow afternoon</T>, the American Red Cross will open **<E>hurricane evacuation centers</E>** for residents who do not feel safe in their homes or live in low-lying areas.

- Entity Extraction
 - Sentence segmentation
 - POS tagging
 - Linear chain conditional random field
 - “BIO” annotation

Information extraction

Service	Time	Status
Miami-Dade County offices	October 24, 2005	close
public schools	October 24, 2005	close
courts	October 24, 2005	close
transit bus	October 22, 2005 6:30 p.m.	open
Rail service	October 22, 2005 6:30 p.m.	open
...		
Miami International Airport	October 22, 2005 6:30 p.m.	open
hurricane evacuation centers	October 23, 2005 afternoon	open

- Relation Extraction
 - (Entity, Time) categorization
 - { No relation, open, close, unclear }

Content recommendation

- Content recommendation
 - Dynamic dashboard
 - Personalized
 - Snapshot
 - Issues to handle
 - Redundancy
 - Rank on relevance

The data sources for different dashboard

Dashboard	EOC Reports	Company Reports	News	Company Messaging
Situation	√		√	
Threat	√			
Event	√	√	√	√
Company		√		√

Content recommendation

- Content Recommendation Engine
 - Content ranking
 - Modeling user preference
 - $I(u)$: set of information submitted by user
 - $J(u)$: set of information viewed by user
 - Scoring the importance

$$f_u = \alpha \frac{\sum_{u \in I(u)} t(u)}{|\sum_{u \in I(u)} t(u)|} + (1 - \alpha) \frac{\sum_{u \in J(u)} t(u)}{|\sum_{u \in J(u)} t(u)|}$$

$$score(d_i) = \text{sim}(f_u, t(d_i)) \cdot pr(d_i)$$

* $t(u)$ - tf-idf of document submitted by u.

* $pr(d)$ - the priority of d.

Reports summarization

- Reports summarization
 - Multiple reports
 - Unstructured/Structured information
 - Sentences: raw, entity/relation identified
 - Time sensitive
 - Documents: within 48 hrs. and beyond 48 hrs.

Reports summarization

- Reports summarization
 - Summarize recent unstructured reports

$$\max_{s_j \in S - V_{k-1}} [I(s_j, l) - \frac{1}{k-1} \sum_{s_i \in V_{k-1}} I(s_j, s_i)],$$

- Summarize recent structured reports

$$\max_{e_j \in E - E_{k-1}} [I(e_j, l) - \frac{1}{k-1} \sum_{e_i \in E_{k-1}} I(e_j, e_i)],$$

- Most recent sentence associated with entity information will be picked.

Reports summarization

- ✓ **miami international airport is open.**
- ✓ **homestead general airport is open from sunrise to sunset.**
- ✓ **homestead airport remains closed.**
- ✓ **tamiami and opa locka airports are open and operational.**
- ✓ **tamiami airport is open and operational.**

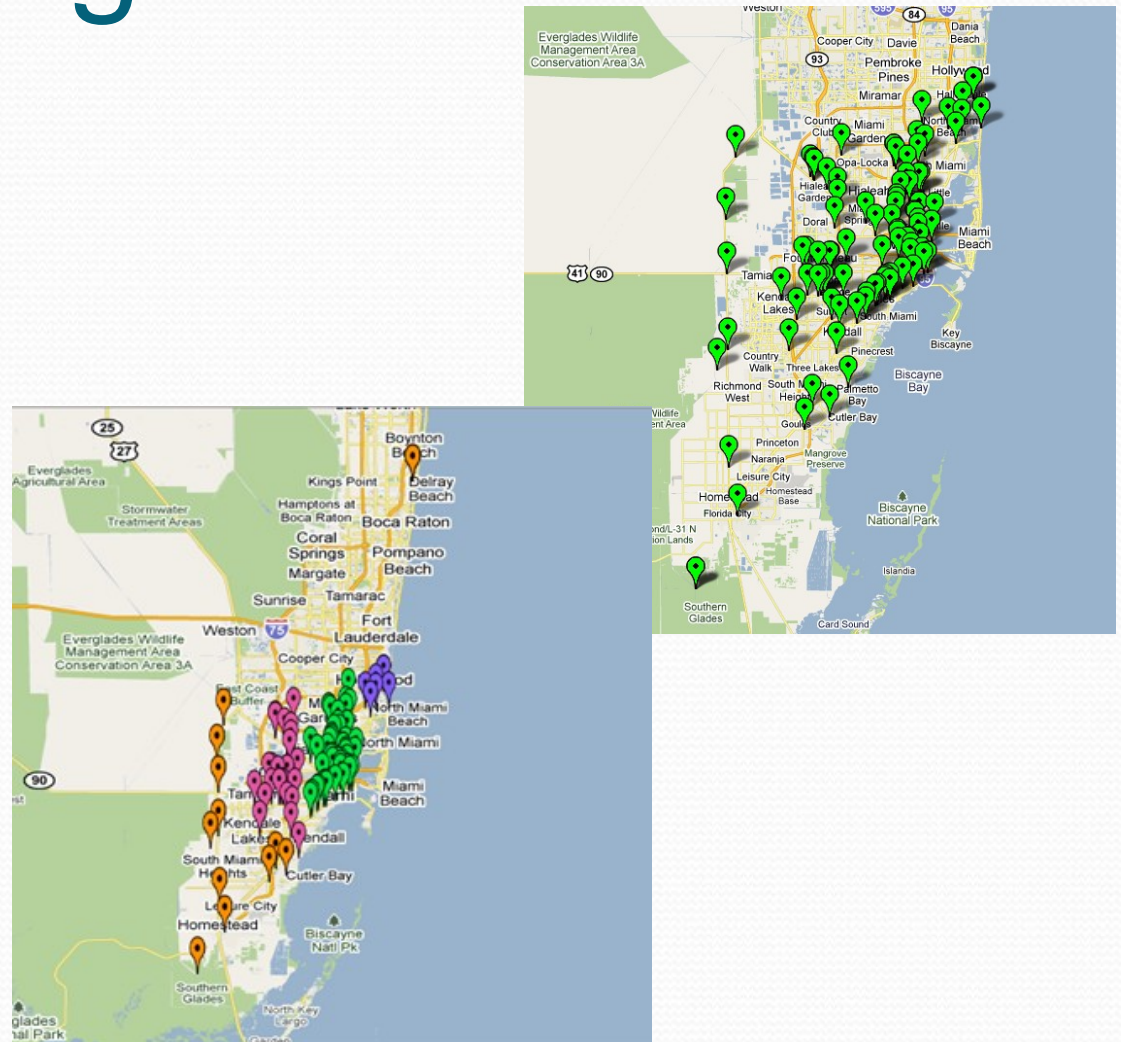
- **Query: Airport**
 - 151 documents
 - October 19, 2005 4:30 p.m. - November 4, 2005 11:30 a.m.

Community generation

- Community
 - Spatial information is associated with every entity
 - Geographical region
 - Similar damage situation
 - Similar recovery status/interests
 - Dynamic community
 - Business relation
 - Status: open/close
 - Region divisions

Community generation

- Constraints
 - Obstacles:
 - Road forbidden
 - Curfew area
 - Facilitators
 - Open expressway



System Exercises

- Exercises for disaster recovery simulation

Date	Description of the Exercise
Jun/01/2009	In Florida Dept. of Emergency Management's Statewide Hurricane Exercises, BCIN was utilized in a scenario where Miami-Dade County Emergency Management Business Recovery Desk facilitated the logistics to deploy portable ATMs at Shelters and PODs in Miami-Dade County.
Jun/29/2009	In Miami-Dade UASI exercise, BCIN was the tool responsible for communicate and collaborating with several companies that participated in the event as observers.
Aug/20/2009	In a full scale company training of BCIN where about 30 companies participated, companies were given injects to provide information to resolve different information requests.

Miami-Dade EM Training Exercise Aug' 2009



Acknowledgements

- This work is supported by NSF grant HRD-0833093 and DHS grant 2009-ST-062-000016. We thank Jason Allen and Mark Oleson for their work in the system development and testing.

- 
- Thank you!