Automatic Crawling of Information to Address the Disaster Information Management

[1]B. Angeline Priyadharshini, [2]G. IlancheziaPandian
1.PG Scholar, Department of Computer Science, GanadipathyTulsi's Jain Engineering College, Vellore, India.
2.HOD- Department of CSE, GanadipathyTulsi's Jain Engineering College, Vellore, India.

Email-id: [1]angelinejaya@gmail.com

Abstract
In recent years, the efficiency of sharing and managing information plays an important role in business recovery efforts after disaster event. The first issue is that reconstructing or creating information flow becomes intractable in domains. The improvement of Crisis Management and Disaster Recovery techniques are national priorities in the wake of man-made and nature inflicted. Users are eager to find valuable information to help them understand the current disaster situation and recovery status.Data mining and information retrieval techniques help impacted communities better understand the current disaster situation and how the community is recovering. Report characterization techniques generate brief reviews from a large collection of reports at different granularity. Probabilistic models with Location specific resource identification algorithm support dynamically generating query forms and information dashboard based on user feedback; and community generation and user recommendation techniques are adapted to help users identify potential contacts for report sharing and community organization.

Keywords: Data Mining, Location Specific Resource Identification Algorithm, Query Forms, Information Retrieval, Recommendation Tehniques.

1.Introduction

The consequence of the interruption in the information flow is that the ability and the efficiency of communication degrade once critical networks are disrupted under disaster impact. Another issues is that need to dissipate a large volume of disaster situational information. It quickly reassemble or create information flow for multi-party coordination activities during disaster situations. The information should be amalgamated from heterogeneous sources. A collaborative platform for preparedness and recovery that helps disaster impacted communities to better understand what the current disaster situation is and how the community is recovering. The intelligent information consignment techniques to help users quickly identify the information.

The deployment of the BCIN, in a disaster situation, coetaneous reports available from thousands of participants would make participants, assess the status without dedicating a significant amount of time by all parties to process this potentially huge volume of information. The system establishes four key capabilities: Messaging, Reporting, Resources, Situational Browsing. BCIN displays user submitted information but also conducts necessary and meaningful data processing work. BCIN makes recommendations based on current focus which dynamically adapts based on the users' interests. BCIN integrates reports for users with brief and content-oriented stories, protects users from trouble while searching in huge amount of information. BCIN offers users a hierarchical view of important reports or events around them. Users share their information by reports, which are mainly about the status of the entities that the users are related with. Reports from County Emergency Management Offices through an Emergency Operation Center, is activated during storm threat, called as EOC reports. Reports from companies are called company reports and messages, the difference of which is that company reports are accessible to users via group and role based access controls while messages are only received by specified target users. It automatically removes the redundant companies' reports, news and other information by clustering methods It ranks the information by both the relevance to the current user and the importance of information.

The logical collection of interdependent individuals, resource values, environmental conditions and related aspects can be identified to form a system. A given universe of discourse (UoD) can be identified to have multiple such systems. Individuals and a group of unique individuals embodied as organizations play specific roles in the UoD. Based on the feigned roles they have to subscribe to appropriate information management strategy that allows identification, processing and delivery of information within the system. To support this information flow, data management strategy must be equipped with appropriate information system.

2.Related Work

An interactive spatial clustering interface for users access multilevel communities in a top–down manner and consider physical or nonphysical obstacles when generating spatial clusters to form more practical communities. It deals the unbalanced size of clusters, we provide users with an interactive mechanism to track the sub community information within a large size community. Clustering process is triggered in the runtime when a user selects a larger community and wants to see the cluster information within such a community at a finer granularity.The model is a best fit for the digital government paradigm as it enables collaboration and communication among major role players for effective disaster

preparedness, recovery measures. To design and implement a web based prototype implementation of our Business Continuity Information Network (BCIN) for rapid disaster recovery system, states hurricane Wilma as the case study.

3.Problem Identification

A critical problem in a crisis situation is how to efficiently devise, convene, codify, search and disseminate real-time disaster information.we have identified several key problems that inhibit better information sharing and collaboration among both private and public sector participants for disaster management and recovery. BCIN utilizes the latest advances in database, data mining, and information extraction technologies to create a user-friendly, The problem with such an approach is that preparedness and response activities do very little to address rising disaster losses. When a disaster occurs, people and resources will flow to the scene and new organizations will appear almost instantaneously. Many of these volunteers are untrained and may create serious problems for first responder. Management theory– Disasters are political and organizational problems. Some of the vulnerability in our communities may be corrected through effective leadership and strategic planning.

4.Proposed System

The primary goal of these systems are message routing, resource tracking, and document management for the purpose to support situation awareness. The proposed technique makes use of iinformation and sharing and collaboration. Two system has been designed and implemented. The system is a web-based prototype of a Business Continuity Information Network system and an All-Hazard Disaster Situation Browser system. Report summarization techniques produce brief views from a large collection of reports at different granularities. Probabilistic models forecast lively emerging query forms and information dashboard based on user feedback.

The features of the proposed system are:

- The user query is received using basic query form.
- The query processing done easily due the query received using basic query form.
- The input data is integrated form different sources easily by making use of report summarization technique.
- An information-rich service on both web-based and mobile platforms in the disaster management domain.
- Hierarchical summarization to automatically extract the status information from a large document set.
- The value that achieves the bestclustering performance
- The different levels ofredundancy and accuracy, possibly generated by a variety of re-ports.

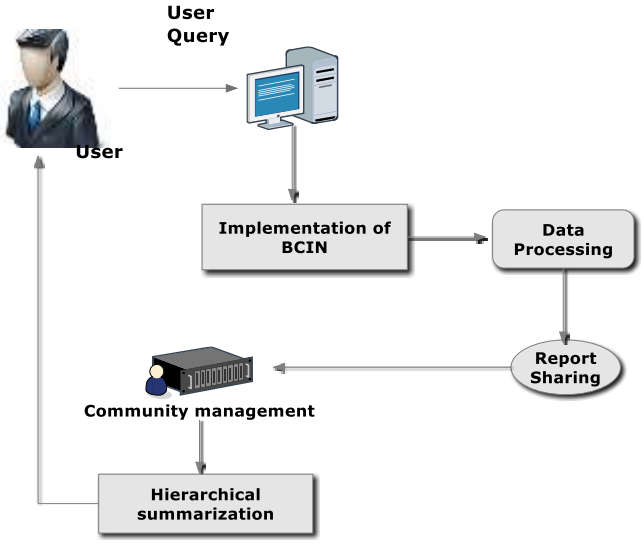

Fig. 1: System Architecture for User making Queries.

A disaster preparation and recovery framework that can be utilized under different disaster conditions and that has the ability to incorporate and utilize multimedia data elements such as videos, audios along with text based input instance plans along with the necessary updates to these programs. The users can then utilize these suggestions to advise employees, other businesses and NGOs etc. about the assistance plans helping them in recovering some portion of the disaster inflicted damages. We also have embedded some level of intelligent decision making regarding preparation and recovery resources by utilizing our local area and location specific resource identification algorithm.

The consequence is that the ability and the efficiency of communication degrade once critical networks are disrupted by the disaster and people may not have alternative paths to transfer information. For example, once power is disabled and uninterruptable power supplies fail after a hurricane, computing and networking equipment will fail unless preventative measures are taken. However, maintaining a fuel-consuming generator is not always possible.

Algorithm

Spatial clustering algorithm

Clustering is a descriptive task that seeks to identify homogeneous groups of objects based on the values of their attributes . In spatial data clustering , those chunks permits a generalization of the spatial component like explicit location and extension of spatial objects which define implicit relations of spatial neighborhood. Current spatial chunk techniques can be broadly classified into three categories; Spatial data clustering identifies chunks, densely populace regions, based on some distance measurement in a large, multidimensional dataset. Many spatial chunk techniques have been developed to identify clusters with arbitrary shapes of various densities and with different physical constraints.

Location specific resource identification Algorithm

The disaster recovery and resources identification methodology thus establishes a virtual marketplace where businesses and emergency management officials can collaborate to expedite the pre-disaster preparation and post-disaster recovery process.

A. Dataset Collection

In this section , we describe the process of collected data. Most commonly a data set dovetail to the contents of a single database table, or a single demographic data matrix, where every column of the table represents a particular variable, and each row dovetail to a given member of the data set in question. The data set provides values for each of the parables, such as height and weight of an object, for each member of the data set. Each appraisal is known as a datum. The data set may encompass data for one or more members, dovetailing to the number of rows.

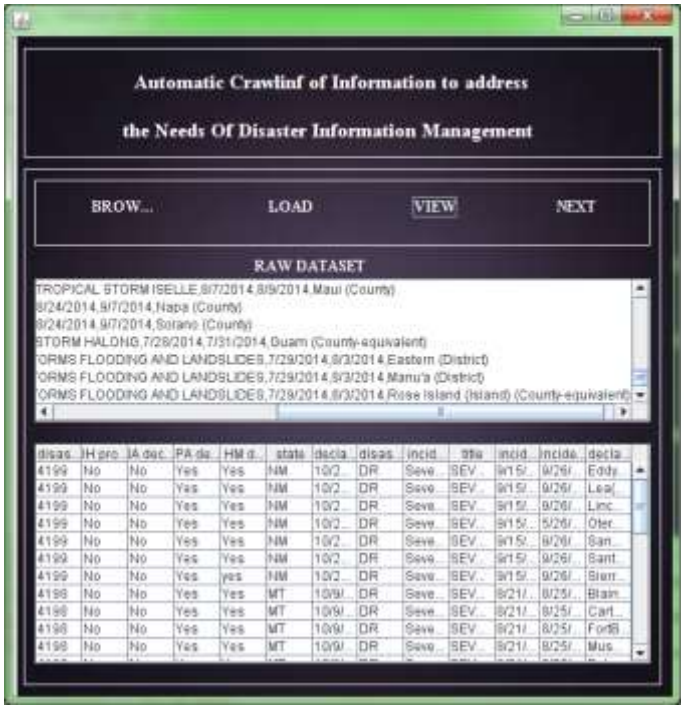

Fig. 2: Collection of Datasets.

B. Preprocessing

After collecting the datasets , Data Preparation and filtering steps can take considerable amount of processing time. Includes cleaning, normalization, transformation, feature extrication and selection etc.Analyzing data that has not been carefully screened for such problems can produce misleading results. Thus, the representation and peculiarity of data is first and foremost before running an analysis. After loading the dataset, the redundant data is removed to get the original data.

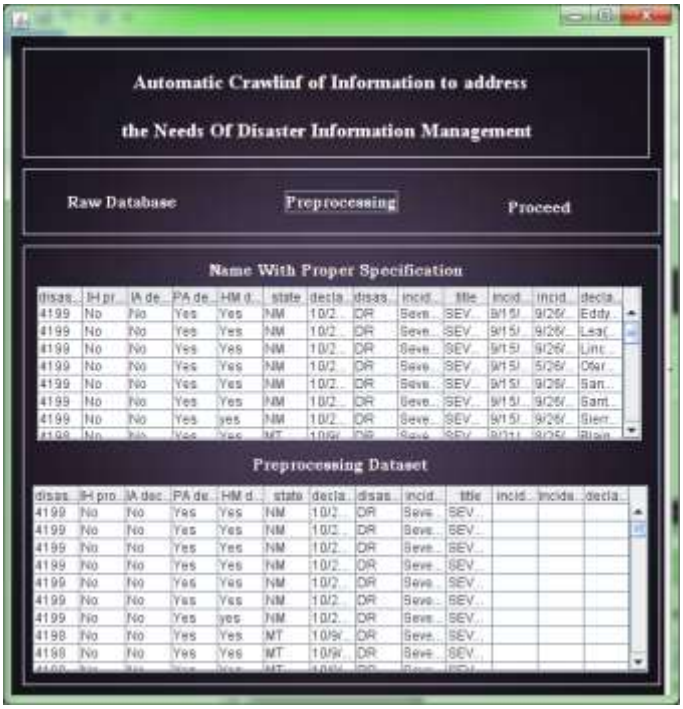

Fig. 3: Preprocessing the datas.

C. Spatial Clustering

In this module Spatial clustering can be used as a standalone tool to gain insight into the distribution of data, to observe the indicative of each cluster, and to focus on a particular set of cluster for analysis. It will operate on the detected cluster. These spatial clustering can be classified into four categories: partitioning method, hierarchical method, quantity based method and grid based method.

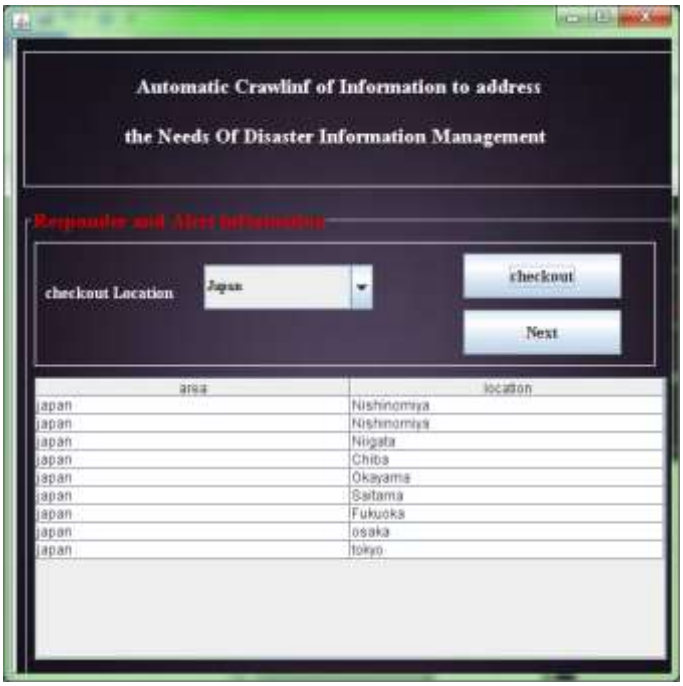

Fig. 4: Spatial Clustering.

Fig. 5: Prediction of datas.

D. Sharing Report

After Spatial clustering process, an interaction is defined as the process of a user sharing a report with multiple users. The report sharing transaction database can be treated as a hyper graph with each node representing a registered user and a set of edges created at the same time from one node to a set of nodes representing an occurred transaction. The efficiency of sharing and management of information plays an important role in the business recovery in a disaster .Users fiind valuable information to help them understand the current disaster situation and recovery status.

E. Dynamic Query Form

A dynamic query form is designed to improve information exploration quality on mobile platforms. It captures users' interests by interactively allowing them to refine and update their queries. To address the fourth challenge, for community discovery, we adopt spatial clustering techniques to track assets like facilities, or equipment, which are important to participants.

7.Conclusion

In this paper, we identified four key design challenges to support multiparty coordination during disaster situations. A unified framework systematically integrates the different techniques that are developed in our previous work. Framework deals with different systems or applications separately and they are necessary collaborative platforms for preparedness and recovery that helps disaster impacted communities to better understand what the current disaster situation is and how the community is recovering. The system evaluation results demonstrate the effectiveness and efficiency of our proposed approaches. The system implementation and assessment process gives the users with suggestions, limitations and possible modifications. A collection of objects and their trait in retrospect with the available physical resources are visualized to form such a system. Extraneous forces act independently in either healing or harming this physical system. The instantaneous state of the system is controlled by the dominating force. Recurrence of sudden unpredictable events may further deviate the system from the desired state.

References

[1] [H. Muson, "Preparing for the worst: A guide to business continuity planning for mid-markets," Executive Action Series, The Conference Board, Rep. A-0179-06-EA, Feb. 2006.

[2] FEMA public private partnership models. [Online]. Available:http://www.fema.gov/privatesector/ppp_models.shtmunderMiami-DadeCounty.

[3] K. Saleem, S. Luis, Y. Deng, S.-C.Chen, V. Hristidis, and T. Li, "Towards a business continuity information network for rapid disaster recovery," in *Proc. Int. Digit. Gov. Res. Conf.*, 2008, pp. 107–116.

[4] V.Hritidis S.chen T Li. S Luis and Y.Deng, "Survey of data management and analysis in disaster situations." J Syst Softw.. vol 83. pp 1701-1714. 2010.

[5] L. Zheng, C. Shen, L. Tang, T. Li, S. Luis, S. Chen, andV. Hristidis, "Using data mining techniques to address critical information exchange needs in disaster affected public-private networks," in Proc.16th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2010, vol. 10, pp. 125–134.

[6] L.Zheng. C. Shen, L. Tang, T. Li, S. Luis and S. Chen "Applying data mining techniques to address disaster information management challenges on mobile devices." in. Proc. Int. Conf. Knowl. Discovery Data Mining, 2011, vol.11,pp.283-291.

[7] D. McEntire, "The status of emergency management theory: Issues, barriers and recommendations for improved scholarship," presented at the FEMA Higher Education Conf., Emmitsburg, MO, USA, 2004.

[8] C.X.Zhai and J.Lafferty, " A study of smoothing methods for languages models applied to ad hoc information retrieval," in Proc. SIGIR, 2001 , pp.334-342.

[9] C. Lin, "Rouge: A package for automatic evaluation of summaries," in Proc. WorkshopACL,2004,pp.25-26.

[10] The Puerto Rico Disaster Decision Support Tool (DDST).[Online].Available:http://www.udel.edu/DRC/DDST/

[11] E. J. Bass, L.A. Baumgart, B. Philips, K. Klossel, K. Dougherty, H. Rodriguez, W. Diaz, W. Donner, J. Santos, and M.Zink, "Incorporating emergency management needs in the development of weather radar networks." J. Emergency Manage., vol. 7, no. 1, pp. 45-52,2009.

[12] L.A.Baumgart, E.J.Bass, B.Philips, and K.Kloesel, "Emergency management decision-making during severe weather", Weather Forecasting, vol. 23, no. 6, pp. 1268-1279,2008.
[13] C. E. League, W. D´ıaz, B. Philips, E. J. Bass, K. A. Kloesel, E. C. Gruntfest, and A. Gessner, "Emergency manager decision-making and tornado warning communication," *Meteorological Appl.*, vol. 17, no. 2, pp. 163–172, 2010.
[14] L. Li and T. Li, "An empirical study of ontology-based multi-document summarization in disaster management," IEEE Trans. SMS:Syst., 2013,in press.