© 2015 IJSRSET | Volume 1 | Issue 2 | Print ISSN : 2395-1990 | Online ISSN : 2394-4099
Themed Section: Engineering and Technology

Ontology Based Recovery of Geographic Information Services

S. Nivedha[1], B. Sneha[2], Sonakshi Bhatoa[3], M. Gayathri[4]
Dhanalakshmi College of Engineering, Chennai, Tamilnadu, India

ABSTRACT

In this paper, we efficiently analyze the trend of the disasters and minimize the consequent loss of data in future to manage expectations, clarity in scope and regular updates. So that false expectation is not created regarding potential use for which the system is not designed. A myriad of news and reports that are related to the disaster may be recorded in the form of text documents. Here we use the concept of Domain Ontology as a meaningful framework for semantic representation of textual information. The domain experts expect to obtain condensed information about the detailed disaster event description. We use multi document summarization technique sentence mapping to summarize multiple documents to get the condensed information and we use KPI algorithm to increase the efficiency to recover document.

Keywords: Disaster management, multi-document summarization, ontology, KPI-key performance indicator.

I. INTRODUCTION

A natural disaster is a sudden calamitous event that causes serious disruption in normal function of the society. Natural disaster is an adverse event which causes destruction of life and property. Any typical disaster will also reflect in economic damage as well. In order to know about the impact of the disaster and minimize the damage in future, we take steps to document all information related to the disaster. This document consist precise information about the origin of the disaster, the intensity and steps taken to overcome the loss caused by the disaster.

A simple example of power outage in Miami is taken as a disaster and disaster information document is prepared by taking into consideration the following details.

1. Date, time and location of the occurrence.
2. Impact of the disaster.
3. Total loss incurred.
4. Steps taken by the officials to recover from disaster.
5. Time taken for recovery.

Power supply in Miami

Florida Power and Light reports 380 000 customers have lost power on October 21th, 2005.

Nearly one million FPL customers without power in Miami-Dade County on October 24th, 2005.

Power is beginning to be restored to FPL customers from October 25th, and it may take several weeks to be fully restored.

The disaster information table normally consists of a list of descriptive sentence that provide the overall summary of the disaster being discussed.

Generally there is data damage during disaster which leads to loss of important and confidential files. Our project is to recover the file lost in a more efficient way by using KPI techniques and various other algorithms. This is further explained in the next section of the paper.

II. METHODS AND MATERIAL

Proposed Idea

In present scenario, if any Disaster happens suddenly, it leads to physical destruction, due to that there is a chance of losing the entire files, the entire database will be deleted and there will be no chance to get the backup data to restore the files in the working database.

In the existing system we use the matching algorithm to support service discovery. We build a prototype as a middle ware that maintains a set of registries (catalog) recording descriptions of geospatial services. But it is very complicated to maintain client's details, Because of power issues.

In order to overcome the defects of existing system, we use Distributed File System Technique where it allows

many clients to have access to the same data/file providing important operations (create, delete, modify, read, and write). Each file may be partitioned into several parts called chunks. Each chunk is stored in remote machines.

The advantages of the above mentioned system is,

1. The recovery from the backup is so fast.
2. Exact file can be retrieved using KPI technique.
3. We can maintain communication for long duration.
4. Cost and complexity is very less compared to Existing System.

Techniques

To implement the proposed system we use specific algorithm which are discussed below:

A. Sentence Mapping

Ontology in disaster management domain provides us abundant conceptual and semantic information, which might facilitate the procedure of multi-document summarization. To utilize the ontology for better understanding the documents, we initially decompose the collection of domain-specific documents into sentences, and then map each sentence to the ontology hierarchy. For each concept of the ontology hierarchy, a group of keywords (i.e., nouns) are assigned by the experts for the sake of sentence mapping. The procedure of sentence mapping is executed based on the following criteria.

1) If the sentence is related to only one concept, map this sentence to the corresponding concept.
2) If the sentence is related to two or more concepts, map this sentence to the least common ancestor (LCA) of these concepts. If the LCA is the most general concept of the ontology, then map the sentence to the original specific concepts.

In this process, we calculate the word set overlapping between a sentence (only considering nouns in the sentence) and the keyword set assigned to each concept as the measure of relatedness, and then rank the scores to select the most related concept. Since different concepts in the ontology have different unambiguous representative noun sets assigned by domain experts, it is unlikely that the same noun will appear in more than one concept. When the condition of the second criterion holds, it means that the sentence contains different words that can map to different concepts. In order to avoid that a single sentence will be linked to multiple concepts and thus make more redundant information, we introduce the LCA of concepts and link the sentence to the LCA if it contains two or more concepts. Based on these criteria, we can guarantee that most sentences are mapped to at least one concept of the ontology because the ontological concepts are representative in a specific domain, and the mapping is reasonable since the mapped sentences can be regarded as instances of the corresponding concept.

B. Key Performance Indicator

Key Performance Indicators (KPIs) can be defined as measures that provide managers with the most important performance information to enable them or their stakeholders to understand the performance level of the organization. KPIs should clearly link to the strategic objectives of the organization and therefore help monitor the execution of the business strategy.

KPI can be defined as providing the most important performance information that enables organizations or their stakeholders to understand whether the organization is on track or not.

KPI can be defined as providing the most important performance information that enables organizations or their stakeholders to understand whether the organization is on track or not.

C. Distributed File System

It is a file system that allows many clients to have access to the same data/file providing important operations (create, delete, modify, read, and write). Each file may be partitioned into several parts called chunks. Each chunk is stored in remote machines.

The distributed file system can be brought in by:

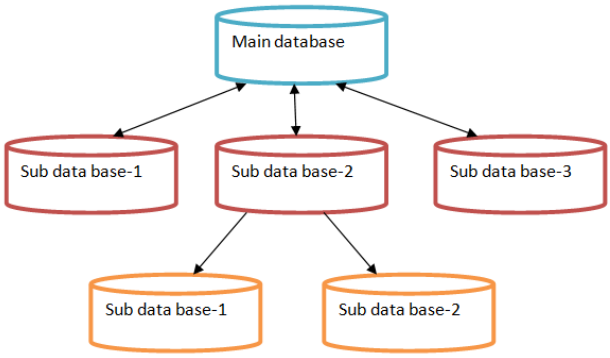

Figure 1: Distributed file system

Distributed file system are essential for sharing of data and storage space in a distributed system. A view point that emphasizes the dispersed structure and decentralization of both data and control in the design of such systems is established. The concepts of location transparency, fault tolerance and scalability are defined and discussed in the context of Distributed File System. It is claimed that the principle of distributed operation is fundamental for a fault tolerant and scalable distributed file system.

III. RESULTS AND DISCUSSION

The implementation begins with the user interface, if existing user he/she has to login, if new user, he/she has to register. After account is created they are allowed to upload files based on disasters. Once they upload, the system generates a key for reference. As soon as they upload, the file is saved in Distributed Data Base and a Backup as a Clustered data base. The key is generated with respect to the file location of existing database and also the back up. The user has to save the key for each and every file manually. If disaster happens the data base should always be available for the user. For that purpose we summarize multiple documents of disaster using sentence mapping to know the intensity of the disaster. Hence using KPI technique we improve the efficiency of the search to get the appropriate document.

The workflow of the system

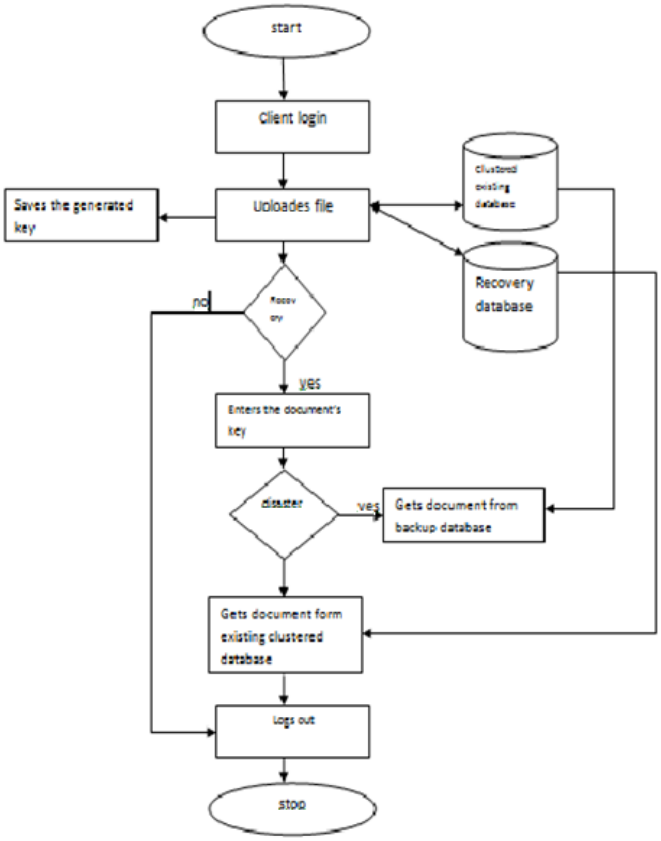

Figure 2: workflow of the system

Some values and outcome of the system:

1. To get the document recovered at any cost of disaster, how intensive it might be.
2. Retrieval of documents at a faster pace.
3. The key technique is used to get the appropriate document effectively.
4. In future, the key database in the client side will be developed so that, the user need not save his document key or to memorize. In such a system, the database will have the name of the document uploaded and also the corresponding key generated.
5. To use the database of key with client login virtually to save the storage space effectively.

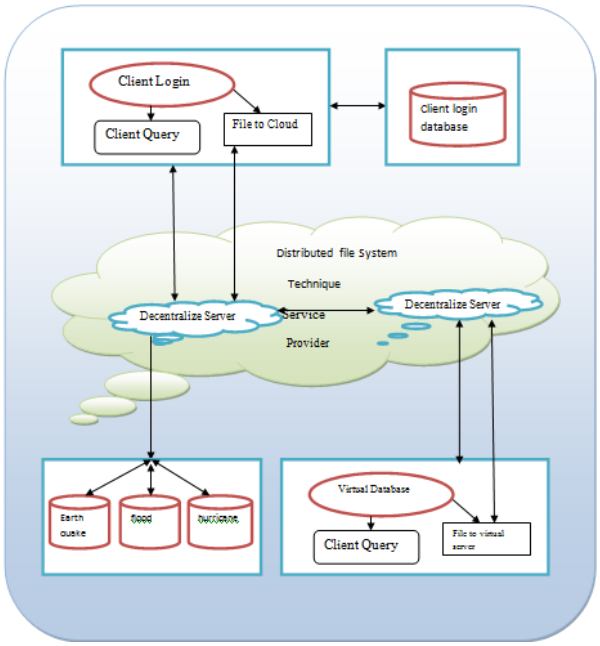

Figure 3: System Architecture

IV. CONCLUSION

This system is formulated to efficiently bring back the existing data base after disaster and it can also be enhance by providing dynamic key storage for the file that is being uploaded in a virtual database. This promises even more efficient and prominent system.

The advantages of the system that is discussed in this paper are:

1. Cheap and efficient.
2. Quick recovery.
3. More appropriate data can be retrieved.
4. File and its data are never unavailable.

V. REFERENCES

[1] D. Radev, H. Jing, M. Sty, and D. Tam, "Centroid-based summarization of multiple documents," Inf. Process. Manage., vol. 40, no. 6, pp. 919–938, 2004.

[2] V. Nastase, "Topic-driven multi-document summarization with encyclopedic knowledge and spreading activation," in Proc. EMNLP, 2008, pp. 763–772.

[3] C. Lee, Z. Jian, and L. Huang, "A fuzzy ontology and its application to news summarization," IEEE Trans. Syst., Man, Cybern., B Cybern., vol. 35, no. 5, pp. 859–880, Oct. 2005.

[4] H. Saggion, K. Bontcheva, and H. Cunningham, "Robust generic and query-based summarisation," in Proc. ECAL, 2003, pp. 235–238.

[5] F. Wei, W. Li, Q. Lu, and Y. He, "Query-sensitive mutual reinforcement chain and its application in query-oriented multi-document summarization," in Proc. SIGIR, 2008, pp. 283–290.

[6] X. Wan, J. Yang, and J. Xiao, "Manifold-ranking based topic focused multi-document summarization," in Proc. IJCAI, 2007, pp. 2903–2908.

[7] O.H. Ibarra and C.E. Kim, "Heuristic Algorithms for Scheduling Independent Tasks on Nonidentical Processors," J. ACM, vol. 24, pp. 280-289, Apr. 1977.

[8] X. Meng et al., "Efficient Resource Provisioning in Compute Clouds via vm Multiplexing," Proc. IEEE Seventh Int'l Conf. Autonomic Computing (ICAC '10), pp. 11-20, 2010.

[9] J. Sonneck and A. Chandra, "Virtual Putty: Reshaping the Physical Footprint of Virtual Machines," Proc. Int'l HotCloud Workshop in Conjunction with USENIX Ann. Technical Conf., 2009.

[10] D. Gupta et al., "Difference Engine: Harnessing Memory Redundancy in Virtual Machines," Proc. Eighth Int'l USENIX Symp. Operating Systems Design and Implementation, pp. 309-322,2008.

[11] O. Sinnen, Task Scheduling for Parallel Systems, Wiley Series on Parallel and Distributed Computing. Wiley-Interscience, 2007.

[12] E. Klien, M. Lutz, and W. Kuhn, "Ontology-based discovery of geographic information services—An application in disaster management," Comput., Environ. Urban Syst., vol. 30, no. 1, pp. 102–123, 2006.

[13] H. Hsu, C. Tsai, M. Chiang, and C. Yang, "Topic generation for web document summarization," in Proc. IEEE SMC, 2008, pp. 3702–3707

[14] X. Yong-dong, W. Xiao-long, L. Tao, and X. Zhi-ming, "Multi-document summarization based on rhetorical structure: Sentence extraction and evaluation," in Proc. IEEE SMC, 2008, pp. 3034–3039.

[15] L. Zheng, C. Shen, L. Tang, T. Li, S. Luis, S. Chen, and V. Hristidis, "Using data mining techniques to address critical information exchange needs in disaster affected public-private networks," in Proc. SIGKDD, 2010, pp. 125–134.