Remote Sensing Data Dissemination and Management: Potential of Replication and Provenance Techniques

Tauqir Ahmad [1], Abad Ali Shah [2]

Department of Computer Science
University of Engineering and Technology, Lahore

tauqir_ahmad@hotmail.com abad_shah@uet.edu.pk

Abstract: With the proliferation of computer technology in almost every sphere of life such as e-government, health care and services sector, there is an increased reliance on digital data. More recently, satellite and remote sensing data has gained importance owning to its applications in real time decisions for GIS, military and strategic needs, and, surveillance and security systems. However, real time access to the digital repositories most of which are web based is plagued by various management and dissemination issues. In data intensive domains such as scientific computations, bioinformatics and e-commerce, replication and provenance have been used successfully for improved performance of data sources by handling the issues of availability, discovery, reliability, authenticity and consistency. In this paper we argue that remote sensing data dissemination and management share common issues and problems with the data intensive domains mentioned above. We also suggest that replication and provenance techniques offer a promising solution to remove the bottleneck of data dissemination and management for real time decision making based on remote sensing data. [Journal of American Science 2010;6(7):188-198]. (ISSN: 1545-1003).

Keywords: Remote sensing data, data dissemination, data management, replication, provenance

1. Introduction

Remote Sensing (RS) is the science of using electromagnetic radiations to identify earth surface features and estimation of their geo-bio physical properties through analysis and interpretation of its spectral, spatial, temporal and polarization signatures using electromagnetic radiations [1]. With the gradual and increased shift of critical and important systems in both private and government sectors towards computerization, remote-sensing data and GIS's are being seen as critical contributor for both short-term and long-term planning in many domains such as social sciences, agriculture, disaster management, security and surveillance, and, military [2]. However, success and reliability of systems based upon remote sensing data and GIS's depends upon its accessibility level and its successful integration with non-spatial data such as socio-economic and population data [2].

The accessibility issues in a remote sensing application consist of remote sensing data acquisition, its maintenance and archiving, and, its dissemination and distribution [3]. The data accessed is then analyzed and interpreted to with other spatial and non spatial data to address various needs. Although, much research is being done on the various stages towards the realization of remote sensing application, the area of remote sensing data dissemination and distribution has not attracted much attention. Although remote sensing data is acquired in real time, its dissemination and distribution constitutes the bottle neck, owning to the various hops over the network in a distributed environment [4]. As remote sensing applications are distributed and real-time, data dissemination and distribution issues form the weakest link in the chain. Interrelated and interdependent to data dissemination and distribution bottlenecks are the issues of remote sensing web service discovery and data provenance. Meaningful and efficient analysis and interpretation of remote sensing data may be significantly facilitated by annotating it with semantic and provenance meta-data [5].

In literature, replication has been used successfully to handle the problems of data dissemination and distribution in data intensive domains such as scientific computations, bioinformatics and e-commerce [4]. Also in [5] authors have proposed a set of factors that may be used as provenance meta-data to enhance the semantic understanding of both web services and web-based data source [6]. However, these techniques have not been widely adapted by remote sensing community to solve the problems of remote sensing data dissemination and management. The reason for this can be traced back to early remote sensing systems that were mostly government owned and centralized [7]. However with the ubiquitous use of remote sensing data in myriad of applications, there has been a gradual shift towards distribution of remote sensing data. This has increased the need for better dissemination and management mechanisms. In this paper we explore the potential of replication and provenance meta-data to enhance the capability of existing remote sensing applications with enhanced data accessibility, service discovery and trust.

The rest of the paper is organized as follows: Section 2 consists of literature survey in which the importance of remote sensing data and its applications is discussed. Also included in this section are the accessibility and dissemination issues of web based data resources and how replication techniques are being used to handle them. Existing work on provenance is also included in this section. In Section 3 we argue the need of replication and provenance techniques for remote sensing data dissemination and management. We end this research paper with conclusions and suggested future work in Section 4.

2. Literature Survey

In the last two decades, increasingly powerful computing and communication technologies have revolutionized the world through the realization of e-systems such as e-governments, e-commerce, and e-sciences [8]. These e-systems are data-driven and data-intensive, collecting myriad of information for both short-term and long-term decision making [8]. The information collected varies from business transactions, medical and personal information to scientific, satellite and surveillance information [9]. The data describing this information may itself be spatial, multimedia or hypertext documents. This diversity of information and data adds an element of complexity to the management and successful execution of e-systems [10]. Conventional database technologies have been deficient and unsuitable to handle this huge amount of data and information stores owning to its diversity in content and format [11]. E-systems have peculiar needs such as effective, efficient and real-time information retrieval [12]. This information further needs to be enriched with semantic information for discovery of patterns. Over the last two decades, most of the data and information resources covering a broad variety of topics have been made available through World Wide Web (WWW) making it the most important repository of data. However, with the unstructured, dynamic and heterogeneous nature of its contents that are inter-connected with hyper-links, WWW adds yet another layer of complexity [13]. In addition to the conventional problems of redundancy and inconsistency, various other issues such as of data provenance, data dissemination, interoperability, and heterogeneity need to be addressed for the successful management of Web-based data and information resources[5],[14].

2.1 Role of Remote Sensing Applications in E-Systems

Amongst the resources that enable the e-systems, remote sensing applications (RSAs) based on spatial and non-spatial data resources have attracted much attention more recently[15]. RSAs are being extensively used for critical applications in various domains such as social sciences, agriculture, disaster management, security and surveillance and, military [16]. In social sciences RSAs are used to monitor environmental conditions, resource management, epidemiology, archaeology and anthropology, international relations, human health conditions, law, and demographic and urban studies [17]. Agriculture benefits from RSAs for example through identification of cultural wastelands, marginal lands and monitoring of temporal behavior of vegetation [18]. Inland wetlands mitigate the harmful effect of pollutants, help control floods, and are used by fish and wild life as breeding, nursery, and feeding grounds. Other direct and indirect benefits include increased yield, improvement of soil fertility, site specific management of agriculture, preservation of eco-diversity and creation of supportive infrastructure such as irrigation [19]. RSAs play key role in effective disaster management. They provide, analyze and interpret data for mapping fire fuels, accessing fire effects, monitoring fire danger, and measuring progress in implementing any fire rescue and control plan [20]. In a similar manner they help design contingency and mitigation plans for electrical outages, floods, tornadoes, earth quacks, volcanic eruption, tsunami, hurricanes, landslides and human caused disaster (e.g. terrorism)[21]. Remote sensing data is also critical to help make timely and intelligent decisions in military operations. Accurate spatial information is critical to the concept of command, control, communication and coordination in military operations. GIS is used by military in a variety of applications that include cartography, intelligence gathering, battle field management, terrain analysis and remote sensing of objects [22].

2.2 Remote Sensing Data: Issues and Problems

In Section 2.1, we discussed the role and importance of remotes sensing data which is used by remote sensing applications for decision making in many E-Systems. However, issues of data dissemination and data provenance need to be addressed if these E-Systems are to make any reliable real time decisions based on remotely sensed data.

2.2.1 Data Dissemination Issues

Remote Sensing Data is difficult to manage and disseminate because of its peculiar characteristics as shown in Table 1 and explained below.

	Remote Sensing Data	**Conventional Data**
i)	Distributed- Temporally and Spatially categorized	Centralized- Static and Stable
ii)	Critical Response time	Not critical;
iii)	Temporal in nature: volatile	Stable: non volatile
iv)	Have validity constraints	Validity is permanent
v)	Evolutionary	Static
vi)	Uncertain, imprecise	Precise, Exact
vii)	Heterogeneous	Homogeneous
viii)	Voluminous	Medium to small scale
ix)	Broad context	Limited Context

Table 1: Comparison of Remote Sensing and Conventional Data

i) Remotely sensed data is either constantly acquired or may be acquired on request for specific purposes. In both the cases, it is organized in some host primary machine using a conventional DBMS [3], [23]. However, fetching fresh data directly from primary site is not efficient. This is because datasets of remotely sensed data may be organized into multiple files containing a subset of data elements categorized temporally and spatially [23].

ii) The response time of such data sets may be adversely affected as multiple queries access these datasets for the same temporal or spatial information [24]. The effect of write operations may even be more adverse as consistency is compromised resulting in conflicting results. Moreover, many real-time applications need to share data that are distributed from a primary site among multiple sites [25].

iii) Remotely sensed data can be categorized into temporal data and non-temporal data. Real time sensor data representing the physical world is temporal in nature [26]. Temporal data objects are difficult to manage and analyze as they have validity intervals and are updated periodically. For example, temperature zones, its spread, location is example of temporal data in a fire management application [27, 28].

iv) As compared to temporal data, non-temporal data do not change dynamically with time. Thus they do not have validity intervals and they do not need to be updated by periodic system updates. However, as most of these data and applications are hosted in a distributed environment, the issues of data dissemination relating, availability, consistency and response time affect non-temporal data also [29].

v) Remote Sensing Data is highly evolutionary, uncertain and incomplete. The latest findings, weather conditions, natural and human activities may invalidates the previous information and data.

vi) Remote sensing data is heterogeneous consisting of sequence data, graphs, high-dimensional data, geometric information, scalar and vector fields, patterns, constraints, images, spatial information, models, prose and declarative knowledge such as hypotheses and evidence

vii) High volume of remote sensing data must be gathered and then disseminated Remote sensing data must have i) ***Factor comprehensiveness***, which reflects the numbers of objects that can be measured at once ii) ***Time-line comprehensiveness***, which represents the time frame within which measurements are made (i.e., the importance of high-level temporal resolution), and iii) ***Item comprehensiveness***—the simultaneous measurement of multiple items, All of these considerations suggest that in addition to being highly heterogeneous, remote sensing data must be voluminous if they are to support comprehensive investigation and interpretation.

viii) The usage of remote sensing data is very broad, with different perceptions, contexts and objectives.

Conventional databases and information technologies have many limitations to overcome the problems mentioned above. As discussed in Section 1, other data-intensive domains such as bioinformatics share the above mentioned issues and problems. In Section 3, we will describe how replication techniques, which have been used to solve similar problems in scientific domains such as bioinformatics, can be applied for remote sensing data dissemination and management.

2.1.2 Data Provenance Issues

With the increased capability and usage of the Internet, various research groups with common interests have moved towards collaborative research [30]. Due to heterogeneity and distribution of data resources, the usability of data resource for a particular domain depends upon the provenance information attached to the data resource. More generally, Data Provenance is the information about data, describing its origin and sequence of tasks (workflows/processes) that were responsible for its transformation; structurally, logically, physically and/or geographically. It provides a qualitative and quantitative metric to analyze the quality and dependability of the data based on consumer trust of the source of creation and the sources that were responsible for modification. Therefore, data provenance has a significant role to play for remote sensing domain in addressing the concerns of trust, quality and copyright [5].
In Geographical Information Systems (GIS), Provenance information includes description of the lineage of the data product including description of the data source, the transformations used to derive it, references to the control information and mathematical transformations of the coordinates used [5]. Recording lineage helps the user of the data product to decide if the data meets the requirements of their application [31]. Spatial Data Transfer Standard (STDS) is one such lineage recording system that helps in using the data product by deciding whether the data meets the requirements of the domain or not. Lineage Information Program (LIP) is for GIS and is used for informational purposes, update stale data, regenerate and compare data. [31] LIP follows a data-oriented provenance technique [31]. Provenance data relates to spatial layers, the basic dataset of GIS, the transformation algorithm and intended use of the data [5], [31]. Semantic Information is not included. To be useful, the data to be transferred must also be meaningful in terms of data content and data quality We feel that if this lineage information is stored and recorded in the machine readable from using Resource Description Framework (RDF) and Ontologies, an efficient automated validation of results is possible with the potential for increased access to and sharing of spatial data, the reduction of information loss in data exchange, the elimination of the duplication of data acquisition, and the increase in the quality and integrity of spatial data. Annotating data from one organization with enough metadata will enable other organizations to use the purposefully and in more meaningful ways [5]. However, there are only a few data provenance systems catering for remote sensing domain which primarily focus data provenance as an issue [5]. In Section 3, we discuss the potential of replication and provenance techniques to solve the problems of remote sensing data dissemination and management.

3 .Potential of Replication and Provenance Techniques

E-Systems are real-time distributed applications and they need real-time data accessibility for prompt and timely decision making. Providing such real-time data accessibility is, however, a challenging task is due to long remote data accessing delays, network hops and inflexible temporal requirements of real-time transactions. Replication is a technique that has been used to ensure the availability of data in a distributed environment to improve the performance and meet the stringent response time requirements of real-time applications [5]. As discussed in Section 2, a real-time data intensive application should be supported by provenance data to establish its authenticity, trust and reliability [5, 13]. More importantly, the provenance data can also be used to avoid duplication of the data replication efforts [5]. As early remote sensing applications were primarily government owned and centralized, the need for replication and provenance techniques was not realized and needed. Examples of such systems are those systems which were used by military for defense, intelligence and reconnaissance purposes [7]. However, with the ubiquitous use of remote sensing (RS) data in myriad of applications such as fire fighting, control surveys, environmental studies, ecology, social sciences etc., there has been a gradual shift towards distribution of RS data [7]. This has increased the need of better dissemination and management mechanisms. We feel that real-time data intensive remote sensing applications can get benefit from the existing replication and provenance techniques. Based on this argument, we advocate the need of applying replication and provenance techniques in remote sensing applications. In our opinion, by applying the replication and provenance techniques, we can address the issues of dissemination of RS data, and achieve the following objectives:

i) Data availability enhancement of data during disk-failure in a multi-disk system,

ii) Speed up I/O performance of read-intensive applications,

iii) Maintain consistency of temporal data,

iv) Maintain reliability of transactions accessing temporal data,

v) Efficient execution of batch-shared I/O oriented data-intensive tasks in web based distributed environment.

RS data is usually "write-once" and "read-many" type of the data and the read performance of an application becomes an important factor [32]. This type of stored data is generally is accessed through range type of queries which retrieve a range of key values from a distributed high dimensional data set [6, 13]. To minimize the disk retrieval time for range queries various non-replication based schemes have been used that distribute data blocks among parallel disks [17, 33]. In these schemes, data files are de-clustered into fragments and spread across multiple disks so that the application can exploit the I/O bandwidth reading and writing the disks in parallel [33]. However, to achieve objectives i) and ii) using the data replication technique, two problems need to be addressed [6, 13, 28 and 29]. These two problems are listed below.

a) The data placement problem of determining the optimal scheme for storing copies of each data block on a particular disk, and

b) the scheduling problem, i.e., to minimize the retrieval time from the disk for a data block.

The replication technique is classified into two categories, i.e., *eager* and *lazy* [6, 11]. In the eager replication approach, all copies of a data item are updated by a single transaction, and his makes it relatively easy to guarantee the consistency as in one-copy serializability for the replicated data, the eager protocols are unsuitable if all replicas are not available [28, 34]. Moreover, it performs poorly if a large number of replicas need to be updated [28]. In the lazy replicated approach, separate transactions are used to update replicas while primary copy is updated through a separate update transaction resulting in a faster response time but compromised consistency [13, 28 and 33]. In case of remote sensing applications and data, most of the transactions are read-only. For a set of requested data blocks, scheduling is better achieved through the lazy approach by retrieving data blocks from the earliest updated replica thus minimizing the retrieval time (better response time) [33, 34].

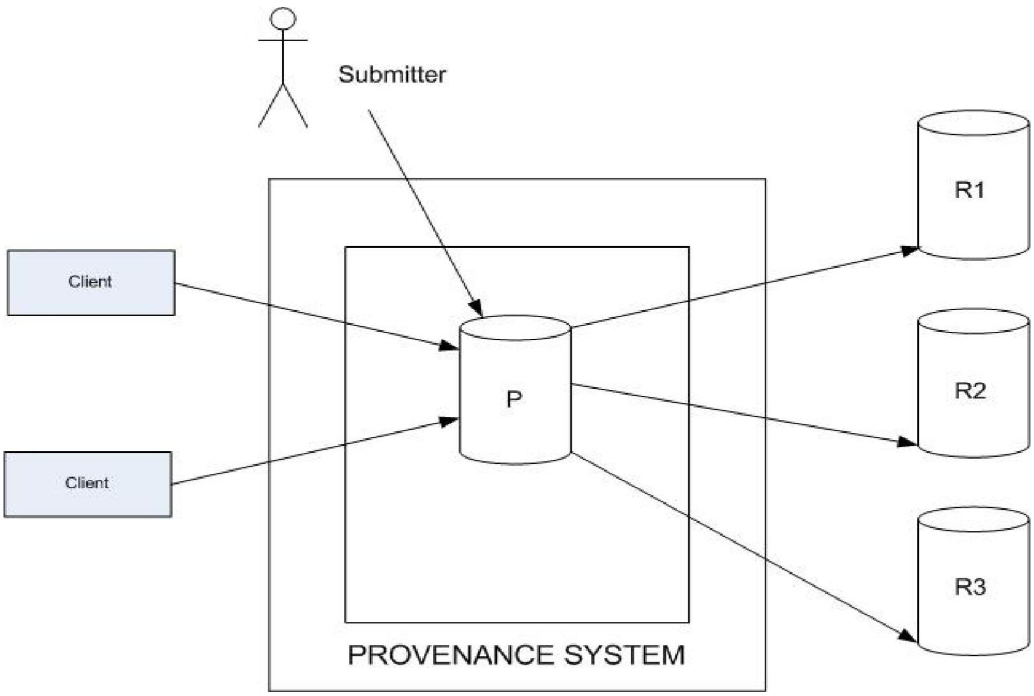
The objectives iii) and iv) may be achieved by hosting both temporal data objects and non-temporal data objects at a primary site which also keeps record of replicas of the temporal data objects hosted within the application [6]. Any recent changes to the temporal data objects, then, may be periodically recorded at the primary site and then updated in the replicas [21]. We are actively working on a hybrid replication model using both the lazy and eager approaches. A lazy approach is a better approach for those applications where most requests are ***read-only*** over temporal data with critical response time requirements [28]. For critical static data, the eager replication approach guarantees consistency [34]. The transactions may be divided into two types: ***system update transactions*** and ***user transactions*** [28]. In the case of remote sensing applications, temporal data update transactions should get priority over replica update transactions to keep consistency [21]. All transactions are validated before accessing fresh data. If the accessed data is stale, then the transaction is restarted. A transaction commits only if it has been validated. We are suggesting that the same validation procedure can be adopted for transactions accessing copies of data available at replicas [11, 12].

Data-Intensive Remote Sensing Applications usually store datasets in collections of files. Such applications may require access to large numbers of files and huge data volume [6, 8]. When mapping tasks to compute nodes, I/O overheads can be minimized by co-coordinating staging of the files because in a batch-shared I/O the same file may be required by multiple tasks in a batch. The objective v) can be achieved by accomplishing the following tasks:

a) to find a mapping of tasks to nodes,

b) to decide which files need to be remotely transferred and their corresponding destination nodes,

c) to determine which files need to be replicated and their corresponding source and destination nodes, so as to minimize the batch execution time.



Legend: **P : Primary Data**
R1, R2, R3: Replicas

Figure 1: Use of replication and provenance for RS data dissemination

In a submarine control system, example of temporal data is its position, speed and depth needed for ship maneuvering. Different combat ships and submarines share this remotely sensed data through real-time distributed database systems. Compared to stand alone local systems, data access is slow in such distributed real-time database systems owning to multi-hop network operations that take substantially more time than the local data accesses. This results in slow response time. Also as a result of delayed transaction completion time, some of the relevant data items used in transaction processing may already have become stale. In a military and defense domain such as a submarine control system, where real time decisions are needed, stale and invalid data may result in disastrous consequences. In this case replication can be used to ensure the availability of data and meet the real time response time requirements. This data needs to be supported by provenance data to establish its authenticity, trust and reliability.

Provenance of data and its replicas enhance the trust, authenticity, and credibility of remote sensing applications [35, 36 and 37]. We suggest that RS data should be annotated with the following provenance meta-data to facilitate its dissemination.

i) **Granularity**: RS data is only useful if it is collected at a required level by the domain and the application. It is for the user to decide the acceptable level of granularity for a particular domain and also recording of replication history.

ii) **Format:** Annotations written in Resource Description Framework (RDF) and Web Ontology Language (OWL) are automatically recorded as they have the advantage of being machine readable.

iii) **Data Core-Elements**: The core elements are: title, description, subject, data and unique identifier which must be recorded.

iv) **Authority**: This parameter determines the trust a user places in a particular data source and its origin.

v) **Timeliness**: This parameter determines how current information is as determined by its current status of usefulness, the age of provenance data (old data tend to be more obsolete) and how often the data source is used.

The provenance data can facilitate the following tasks [39]:

i) Ranking of a data resource,
The user of a web resource may assign a comparative value to it tools reflecting the relevance of the web resource based on its experience.

ii) Updating data resources and its replicas.
Meta-data relating representing the last update to a replica or a resource may be annotated by meta-data relating to its most recent usage

iii) Recommending a data resource
To establish the relevance of a web resource to a user based on usage patterns and user profiles.

iv) Measuring the usefulness of data resources
Calculating how useful a web resource has been to a certain community.

v) Standardizing selection based on a minimum acceptable standard,
A quality analysis using the provenance information can be carried out to determine whether the record complies with a minimum standard to be accepted. This could be especially important for automatic generators of provenance information.

vi) Selection of the most appropriate replication instance [40].
A web resource can have several provenance instances that describe it. The quality could be taken in account when selecting the most appropriate record for each situation (search, evaluation, assembly, etc).

In Figure 1, the submitter which may be a RS data generation /collection entity submits data to a primary site. The submitted data is annotated with provenance meta-data such as submitter authority, data-core elements, format etc., as mentioned above. The data is replicated to R1, R2 and R3 with necessary information. The clients accessing the primary site or its replicas also record necessary information relating to their usage patterns, accessibility and recommendation etc. In this way all transactions are annotated with the necessary meta-data relating to a particular transaction. This annotated meta-data also helps avoid duplication of replication efforts for RS data, remote sensing services and workflows [38].

4 Conclusion and Future Work

Replication and provenance techniques have been used successfully to handle the problems of data dissemination in Web-based data intensive domains such as e-commerce and bioinformatics. However, these techniques have not been adapted by mainstream remote sensing community. With the increasing reliance of E-systems on remote sensing data and remote sensing applications, we feel that replication and provenance techniques must be adapted by remote sensing community if it is to handle the data dissemination issues and problems. In this paper we have proposed that usage of existing replication and provenance techniques to provide a basis for remote sensing data and applications. Although our work is in preliminary stage, we are working on a hybrid model consisting of lazy and eager replication techniques which can cater for peculiar problems of remote sensing data and applications such as its hyper-rectangular nature, spatiality and temporal volatility. We are also extending the list of necessary meta-data categories presented in this paper and working on a provenance system that records and annotates semantic information relating to a data resource and its replica. In our opinion, existing replication and provenance systems in other scientific domains such as bioinformatics will provide us the necessary impetus to realize a functional model of using these techniques for remote sensing applications.

ACKNOWLEDGEMENT
We acknowledge the support extended by Higher Education Commission (HEC), Pakistan and University of Engineering and Technology, Lahore**.**

REFERENCES

[1] Ranganath, R., Navalgund, V., Jayaraman, and Roy, P. S., "Remote sensing applications: An overview", Current Science, VOL. 93, NO. 12, pp. 1747-1766, Space Applications Centre, Ambawadi, India,2007.

[2] Charles, G., David, K, "Land Use Planning and Security in Terrains of Terror", Conference on "the Science and Education of Land Use: A transatlantic, multidisciplinary and comparative approach" in Washington, DC, USA. Sep, 24-26, 2007.

[3] Rajitha, K. , Mukherjee, C.K., and Vinu Chandran, R., "Applications of remote sensing and GIS for sustainable management of shrimp culture in India", Aquacultural Engineering, ScienceDirect, www.elsevier.com/locate/aqua-online, volume 36,No.1, pp. 1–17,USA, 2007

[4] NICHOLAS, M.S., "The Remote Sensing Tutorial". 2007 http://rst.gsfc.nasa.gov/Front/tofc.html

[5] Ahsan, S., Shah, A , "Designing software-intensive systems: methods and principles", Quality Metrics for data provenance, chap. 25.. Published by Idea Group Inc (IGI),USA, pp. 455-473.

[6] Baker, M., Buyya, R. and Laforenza, D.," Grids and Grid technologies for wide-area distributed computing", Software- Practice And Experience (DOI: 10.1002/spe.488) Copyright 2002 JohnWiley& Sons, Ltd. 2002.

[7] John, R., Jensen, "Remote Sensing of the Environment, An Earth Resource Perspective", Pearson Education Singapore, 2000, Ch-3, pp.53-80

[8] Bagchi, S., Kar, G., Hellerstein, J.," Dependency Analysis in Distributed Systems using Fault Injection: Application to Problem Determination in an e-commerce Environment", 12th international workshop on distributed systems: operations and management DSO '2001 Nancy France, 2001.

[9] Fang, Z., "E-Government in Digital Era: Concept, Practice, and Development School of Public Administration", National Institute of Development Administration (NIDA), Thailand, International Journal of The Computer, The Internet and Management, Vol. 10, No.2, , pp. 1-22, Thailand, 2002.

[10] Longley, P., Michael F., David, J.M., David, W.R.," Geographical Information Systems and Science", 2nd edition, Published by JohnWiley and Sons, USA, pp. 3-7, 2005.

[11] PATRICK, V., "Parallel database systems: Open problems and new issues", Distributed and Parallel Databases, Springer. Volume 1, No. 2, pp. 137-165, Netherlands, 1993.

[12] Brown, P. J., Jones, G. J. F., "Context-aware Retrieval: Exploring a New Environment for Information Retrieval and Information Filtering, Personal and Ubiquitous Computing", Publisher, Springer, Volume 5, No. 4, pp. 253-263, London, 2001.

[13] Cooley, R., Mobasher, B. and Srivastava, J.,"WebMining: Information and Pattern Discovery on theWorldWideWeb" Proceedings, Ninth IEEE International Conference on, Tools with Artificial Intelligence, Location: Newport Beach, CA, USA, References Cited: 56, pp. 558-567, 2002.

[14] Goble, C., David, D.R.," The Grid: An Application of the Semantic Web" Volume 31, No. 4, Pages: 65 – 70, ACM SIGMOD Record, USA, , 2002.

[15] Anselin, L., Kim, Y.W., and Syabri, I.,"Web-based analytical tools for the exploration of spatial data", Journal of Geographical Systems, 6 : pp. 197–218, USA, 2004.

[16] Sherbinin, A.D., Deborah, B., Karina, Y., Malanding, J., Francesca, P., Chandra, G. and Antroinette, W.," A CIESIN Thematic Guide to Social Science Applications of Remote Sensing". Chapter 3 updated January 2006 Center for International Earth Science Information Network (CIESIN) Columbia University, 2002.

[17] Ronald, R., Paul, C., stern, R.," Linking remote sensing and social science: the need and the challenges", People and pixels, By Diana M. L., Published by National Academies Press, National Research Council (U.S.), 1998.

[18] Srinivasa, S.R., Murthey, K., Adiga, S. and Emmineedu, E.," Performance Index forWatershed development", J.Ind. Geophys, Union, Vol, No.4, pp. 239-247, India, 2003.

[19] Bambaradeniya, C., "Freshwater Wetlands in Sri Lanka: Their Conservation Significance and Current Status In: IUCN Sri Lanka", Proceedings of the National Symposium on Wetland Conservation and Management, pp. 19-24, 2004.

[20] Emilio, C., Pilar, M., and Chris, J.," Innovative Concepts and Methods in Fire Danger Estimation", 4th International Workshop on Remote Sensing and GIS Applications to Forest Fire Management, Ghent – Belgium, 5-7 June 2003.

[21] Haddi, A.A.El, Lee, V., Banks, J., Sustman, P.," Using File replication for business continuance and content distribution", Accessed on: 2007, Constant Data Inc (www.constantdata.com), (2003).

[22] Stanley, D. Brunn, Susan, L. Cutter, James, W. Harrington, Geography and technology. Published by Springer, USA, ch-18, pp. 401-430, 2004.

[23] Gary, L.G., Cort, J.W.," Geography in America at the dawn of the 21st century", Published by Oxford University Press, Ch, 29. PP.376-418, 2006.

[24] Buetow1, T., Chaboya, L., Christopher, O.T., Cushna, T., Daspit, D. Petersen, T., Atabakhsh, H., and Chen, H.," A Spatio Temporal Visualizer for Law Enforcement", Lecture Notes in Computer Science, Springer-Verlag Berlin Heidelberg, volume: 2665/2003 pp. 181–194, 2003.

[25] Hoschek, W., Javier, JM., Samar, A. Stockinger, H., and Stockinge, K., "Data Management in an International Data Grid Project", Lecture Notes in computer science, Volume 1971/2000, pp. 333-361,. Springer-Verlag Berlin Heidelberg, 2000.

[26] Birant, D., Kut, A.," ST-DBSCAN: An algorithm for clustering spatial–temporal data", Data & Knowledge Engineering, Volume 60, pp. 208–221, Holand, 2007.

[27] Kao, B., Kam, Y.L., Cheng, R.," Maintaining Temporal Consistency of Discrete Objects in Soft Real-Time Database Systems", IEEE TRANSACTIONS ON COMPUTERS, ACM USA, VOL. 52, NO. 3, MARCH 2003

[28] Yuan, W., Andrew A. Aslinger, S.H. John, S., Stankovic, A.," ORDER: A Dynamic Replication Algorithm for Periodic Transactions in Distributed Real-Time Databases", In Proceedings of Real-time and Embedded Computing Systems and Applications (RTCSA04) USA. pp. 152-169, 2004.

[29] Vestal, S., "Real-Time Sampled Signal Flows through Asynchronous Distributed Systems", Proceedings of the 11th IEEE Real Time and Embedded Technology and Applications Symposium (RTAS'05), ISSN: 1080-1812, pp.170-179, USA, 2005.

[30] Foster, I., Kesselman, C., Tuecke, S., "The Anatomy of the Grid: Enabling Scalable Virtual Organizations", International Journal of High Performance Computing Applications, 15 (3). pp. 200-222, SA, 2001.

[31] Yogesh, L., Simmhan, Plale, B. and Dennis, G.," A Survey of Data Provenance in e-Science", ACM SIGMOD Record, Volume 34, No 3 pp. 31–36, USA, 2005.

[32] J.Gray, P.Helland, P.E.O., Neil, and D. Shasha," The dangers of replication and a solution. In *Proc. SIGMOD*, USA, pages 173–182, 1996.

[33] Li Jianzhong, S. Jaideep and Srivastava Doron Rotem, "CMD: A Multidimensional Declustering Method" In Proceedings of the Int. Conf. on Very Large Data Bases, 3–14, 1992.

[34] Sang H. S., and Fengjie, Z.,"Real-Time Replication Control for Distributed Database Systems: Algorithms and Their Performance", Proceedings of the Fourth International Conference on Database Systems for Advanced Applications (DASFAA '95) Ed., Tok Wang Ling and Yoshifumi Masunaga, Singapore, April 10-13, 1995.

[35] Zhao, J., Wroe, C., Carole, G., Stevens, R., Quan, D., and Greenwood, M., "Using Semantic Web Technologies for Representing E-science Provenance", Lecture notes in computer science., Volume 3298/2004 pp. 92–106, 2004. ISSN: 0302-9743 (Print) 1611-3349 (Online), Publisher: Springer Berlin / Heidelberg DOI 10.1007/b102467.

[36] C. Goble. (2002). Position Statement: Musings on Provenance, Workflow and (Semantic Web) Annotations for Bioinformatics,. Workshop on Data Derivation and Provenance, Chicago

[37] D. P. Lanter.(1990). Lineage in GIS: The Problem and a Solution.Technical Report: National Center for Geographic Information and Analysis,.

[38] D. P. Lanter.(1993). A Lineage Meta-Database Approach Towards Spatial Analytic Database Optimization. Cartography and Geographic Information Systems, vol. 20, pp. 112-121.

[39] H.~Veregin and D. P. Lanter.(1995). Data-quality enhancement techniques in layer-based geographic information systems. Computers, Environment and Urban Systems, vol. 19, pp. 23-36

[40] J. Zhao, C. A. Goble, R. Stevens, and S. Bechhofer.(2004). Semantically Linking and Browsing Provenance Logs for E-science. ICSNW, pp. 158-176.

4/1/2010