



www.ijatir.org

Scheming High Performance Web-Based Computing Services to Promote Telemedicine Database Administration System

GRANDHI LOKESH¹, METLAPALLI AYYAPPA CHAKRAVARTHI²

¹PG Scholar, Dept of CSE, G.V.R&S college of Engineering & Technology, Budampadu, Guntur, AP, India,
E-mail: lokesh571@gmail.com.

²Associate Professor, Dept of CSE, G.V.R&S college of Engineering & Technology, Budampadu, Guntur, AP, India,
E-mail: ayyappachakravarthi@gmail.com.

Abstract: Many web computing systems are running real time database services where their information change continuously and expand incrementally. In this context, web data services have a major role and draw significant improvements in monitoring and controlling the information truthfulness and data propagation. Currently, web telemedicine database services are of central importance to distributed systems. However, the increasing complexity and the rapid growth of the real world healthcare challenging applications make it hard to induce the database administrative staff. In this paper, we build an integrated web data services that satisfy fast response time for large scale Tele-health database management systems. Our focus will be on database management with application scenarios in dynamic telemedicine systems to increase care admissions and decrease care difficulties such as distance, travel, and time limitations. We propose three-fold approach based on data fragmentation, database websites clustering and intelligent data distribution. This approach reduces the amount of data migrated between websites during applications' execution; achieves cost-effective communications during applications' processing and improves applications' response time and throughput. The proposed approach is validated internally by measuring the impact of using our computing services' techniques on various performance features like communications cost, response time, and throughput. The external validation is achieved by comparing the performance of our approach to that of other techniques in the literature. The results show that our integrated approach significantly improves the performance of web database systems and outperforms its counterparts.

Keywords: Web Telemedicine Database Systems (WTDs), Database Fragmentation, Data Distribution, Sites Clustering.

I. INTRODUCTION

The rapid growth and continuous change of the real world software applications have provoked researchers to propose several computing services' techniques to achieve more efficient and effective management of web telemedicine database systems (WTDS). Significant research progress has been made in the past few years to

improve WTDS performance. In particular, databases as a critical component of these systems have attracted many researchers. The web plays an important role in enabling healthcare services like telemedicine to serve inaccessible areas where there are few medical resources. It offers an easy and global access to patients' data without having to interact with them in person and it provides fast channels to consult specialists in emergency situations. Different kinds of patient's information such as ECG, temperature, and heart rate need to be accessed by means of various client devices in heterogeneous communications environments. WTDS enable high quality continuous delivery of patient's information wherever and whenever needed. Several benefits can be achieved by using web telemedicine services including: medical consultation delivery, transportation cost savings, data storage savings, and mobile applications support that overcome obstacles related to the performance (e.g., bandwidth, battery life, and storage), security (e.g., privacy, and reliability), and environment (e.g., scalability, heterogeneity, and availability). The objectives of such services are to:

- develop large applications that scale as the scope and workload increases,
- achieve precise control and monitoring on medical data to generate high telemedicine database system performance,
- provide large data archive of medical data records, accurate decision support systems, and trusted event-based notifications in typical clinical centers.

Recently, many researchers have focused on designing web medical database management systems that satisfy certain performance levels. Existing methods consider the three-fold services together which makes them impracticable in the field of web data-base systems. Additionally, using multiple medical services from different web database providers may not fit the needs for improving the telemedicine database system performance. Furthermore, the services from different web data-base providers may not be compatible or in some cases it may increase the processing time because of the constraints on the network. Finally, there has been lack in the tools that support the design, analysis

and cost-effective deployments of web telemedicine database systems. Designing and developing fast, efficient, and reliable incorporated techniques that can handle huge number of medical transactions on large number of web healthcare sites in near optimal polynomial time are key challenges in the area of WTDS. Data fragmentation, websites clustering, and data allocation are the main components of the WTDS that continue to create great research challenges as their current best near optimal solutions are all NP-Complete. To improve the performance of medical distributed data-base systems, we incorporate data fragmentation, websites clustering, and data distribution computing services together in a new web telemedicine database system approach. This new approach intends to decrease data communication, increase system throughput, reliability, and data availability.

The decomposition of web telemedicine database relations into disjoint fragments allows database transactions to be executed concurrently and hence minimizes the total response time. Fragmentation typically increases the level of concurrency and, therefore, the system throughput. The benefits of generating telemedicine disjoint fragments cannot be deemed unless distributing these fragments over the websites, so that they reduce communication cost of database transactions. Database disjoint fragments are initially distributed over logical clusters (a group of websites that satisfy a certain physical property, e.g., communications cost). Distributing database disjoint fragments to clusters where a benefit allocation is achieved, rather than allocating the fragments to all web-sites, have an important impact on database system throughput. This type of distribution reduces the number of communications required for query processing in terms of retrieval and update transactions; it has always a significant impact on the web telemedicine database system performance. Moreover, distributing disjoint fragments among the websites where it is needed most, improves database system performance by minimizing the data transferred and accessed during the execution time, reducing the storage overheads, and increasing availability and reliability as multiple copies of the same data are allocated.

Database partitioning techniques aim at improving database systems throughput by reducing the amount of irrelevant data packets (fragments) to be accessed and transferred among different websites. However, data fragmentation raises some difficulties; particularly when web telemedicine database applications have contradictory requirements that avert breakdown of the relation into mutually exclusive fragments. Those applications whose views are defined on more than one fragment may suffer performance ruin. In this case, it might be necessary to retrieve data from two or more fragments and take their join, which is costly. Data fragmentation technique describes how each fragment is derived from the database global relations. Three main classes of data fragmentation have been discussed in the literature; horizontal, vertical and hybrid. Although there are various schemes describing data

partitioning, few are known for the efficiency of their algorithms and the validity of their results. The Clustering technique identifies groups of network sites in large web database systems and discovers better data distributions among them. This technique is considered to be an efficient method that has a major role in reducing the amount of transferred and accessed data during processing database transactions. Accordingly, clustering techniques help in eliminating the extra communications costs between websites and thus enhances distributed data-base systems performance. However, the assumptions on the web communications and the restrictions on the number of network sites, make clustering solutions impractical. Moreover, some constraints about network connectivity and transactions processing time bound the applicability of the proposed solutions to small number of clusters.

Data distribution describes the way of allocating the disjoint fragments among the web clusters and their respective sites of the database system. This process addresses the assignment of each data fragment to the distributed database websites. Data distribution related techniques aim at improving distributed database systems performance. This can be accomplished by reducing the number of database fragments that are transferred and accessed during the execution time. Additionally, Data distribution techniques attempt to increase data availability, elevate database reliability, and reduce storage overhead. However, the restrictions on database retrieval and update frequencies in some data allocation methods may negatively affect the fragments distribution over the websites. In this work, we address the previous drawbacks and propose a three-fold approach that manages the computing web services that are required to promote telemedicine database system performance. The main contributions are: Develop a fragmentation computing service technique by splitting telemedicine database relations into small disjoint fragments. This technique generates the minimum number of disjoint fragments that would be allocated to the web servers in the data distribution phase. This in turn reduces the data transferred and accessed through different websites and accordingly reduces the communications cost.

Introduce a high speed clustering service technique that groups the web telemedicine database sites into sets of clusters according to their communications cost. This helps in grouping the websites that are more suitable to be in one cluster to minimize data allocation operations, which in turn helps to avoid allocating redundant data. Propose a new computing service technique for tele-medicine data allocation and redistribution services based on transactions' processing cost functions. These functions guarantee the minimum communications cost among websites and hence accomplish better data distribution compared to allocating data to all websites evenly. Develop a user-friendly experimental tool to perform services of telemedicine data fragmentation, web-sites clustering, and fragments allocation, as well as assist database administrators in measuring WTDS performance.

Scheming High Performance Web-Based Computing Services to Promote Telemedicine Database Administration System

II. RELATED WORK

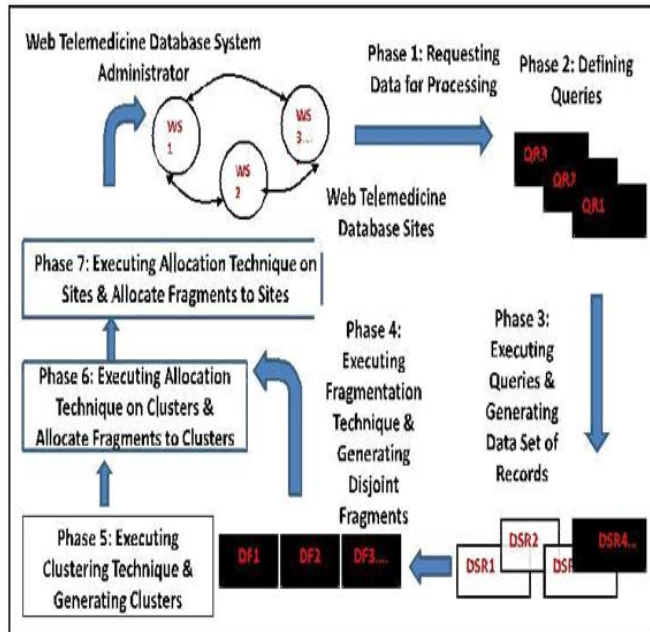


Fig.1. IFCA computing services architecture.

Integrate telemedicine database fragmentation, web-sites clustering, and data fragments allocation into one scenario to accomplish ultimate web telemedicine system throughput in terms of concurrency, reliability, and data availability. We call this scenario Integrated-Fragmentation-Clustering-Allocation (IFCA) approach. Fig. 1 depicts the architecture of the pro-posed telemedicine IFCA approach. In Fig. 1, the data request is initiated from the telemedicine database system sites. The requested data is defined as SQL queries that are executed on the database relations to generate data set records. Some of these data records may be overlapped or even redundant, which increase the I/O transactions' processing time and so the system communications overhead.

To solve this problem, we execute the proposed fragmentation technique which generates telemedicine disjoint fragments that represent the minimum number of data records. The web telemedicine database sites are grouped into clusters by using our clustering service technique in a phase prior to data allocation. The purpose of this clustering is to reduce the communications cost needed for data allocation. Accordingly, the proposed allocation service technique is applied to allocate the generated disjoint fragments at the clusters that show positive benefit allocation. Then the fragments are allocated to the sites within the selected clusters. Database administrator is responsible for recovering any site failure in the WTDS. The remainder of the paper is organized as follows. Section 22 summarizes the related work. Basic concepts of the web telemedicine database settings and assumptions are discussed in Section 3. Telemedicine computation services and estimation model are discussed in Section 4. Experimental results and performance evaluation are presented in Section 35. Finally, in Section 6, we draw conclusions and outline the future work.

Many research works have attempted to improve the performance of distributed database systems. These works have mostly investigated fragmentation, allocation and sometimes clustering problems. In this section, we present the main contributions related to these problems, discuss and compare their contributions with our pro-posed solutions.

A. Data Fragmentation

With respect to fragmentation, the unit of data distribution is a vital issue. A relation is not appropriate for distribution as application views are usually subsets of relations [31]. Therefore, the locality of applications' accesses is defined on the derivative relations subsets. Hence it is important to divide the relation into smaller data fragments and consider it for distribution over the network sites. The authors in [8] considered each record in each database relation as a disjoint fragment that is subject for allocation in a distributed database sites. However, large number of database fragments is generated in this method, causing a high communication cost for transmitting and processing the fragments. In contrast to this approach, the authors in [11] considered the whole relation as a fragment, not all the records of the fragment have to be retrieved or updated, and a selectivity matrix that indicates the percentage of accessing a fragment by a transaction is proposed. However, this research suffers from data redundancy and fragments overlapping.

B. Clustering Websites

Clustering service technique identifies groups of networking sites and discovers interesting distributions among large web database systems. This technique is considered as an efficient method that has a major role in reducing transferred and accessed data during transactions processing [9]. Moreover, grouping distributed net-work sites into clusters helps to eliminate the extra communication costs between the sites and then enhances the distributed database system performance by minimizing the communication costs required for processing the transactions at run time. In a web database system environment where the number of sites has expanded tremendously and amount of data has increased enormously, the sites are required to manage these data and should allow data transparency to the users of the database. Moreover, to have a reliable data-base system, the transactions should be executed very fast in a flexible load balancing database environment. When the number of sites in a web database system increases to a large scale, the problem of supporting high system performance with consistency and availability constraints becomes crucial. Different techniques could be developed for this purpose; one of them is websites clustering. Grouping websites into clusters reduces communications cost and then enhances the performance of the web database system.

However, clustering network sites is still an open problem and the optimal solution to this problem is NP-Complete [12]. Moreover, in case of a complex network where large numbers of sites are connected to each other, a huge number of communications are required, which

increases the system load and degrades its performance. The authors in [13] have proposed a hierarchical clustering algorithm that uses similarity upper approximation derived from a tolerance (similarity) relation and based on rough set theory that does not require any prior information about the data. The presented approach results in rough clusters in which an object is a member of more than one cluster. Rough clustering can help researchers to discover multiple needs and interests in a session by looking at the multiple clusters that a session belongs to. However, in order to carry out rough clustering, two additional requirements, namely, an ordered value set of each attribute and a distance measure for clustering need to be specified [14]. Clustering coefficients are needed in many approaches in order to quantify the structural network properties. In [15], the authors proposed higher order clustering coefficients defined as probabilities that determine the shortest distance between any two nearest neighbors of a certain node when neglecting all paths crossing this node. The outcomes of this method declare that the average shortest distance in the node's neighborhood is smaller than all network distances. However, independent constant values and natural logarithm function are used in the shortest distance approximation function to determine the clustering mechanism, which results in generating small number of clusters.

C. Data Allocation (Distribution)

Data allocation describes the way of distributing the database fragments among the clusters and their respective sites in distributed database systems. This process addresses the assignment of network node(s) to each fragment [8]. However, finding an optimal data allocation is NP-complete problem. Distributing data fragments among database websites improves database system performance by minimizing the data transferred and accessed during execution, reducing the storage over-head, and increasing availability and reliability where multiple copies of the same data are allocated. Many data allocation algorithms are described in the literature. The efficiency of these algorithms is measured in term of response time the author has addressed the fragment allocation problem in web database systems. He presented an integer programming formulations for the non-redundant version of the fragment allocation problem. This formulation is extended to address problems, which have both storage and processing capacity constraints. In this method, the constraints essentially state that there has been exactly one copy of a fragment across all sites, which increase the risk of data inconsistency and unavailability in case of any site failure. However, the fragment size is not addressed while the storage capacity constraint is one of the major objectives of this approach. In addition, the retrieval and update frequencies are not considered in the formulations, they are assumed to be the same, which affects the fragments distribution over the sites. Moreover, this research is limited by the fact that none of the approaches presented have been implemented and tested on a real web database system.

A dynamic method for data fragmentation, allocation, and replication is proposed in [25]. The objective of this approach is to minimize the cost of access, re-fragmentation, and reallocation. DYFRAM algorithm of this method examines accesses for each replica and evaluates possible re-fragmentations and reallocations based on recent history. The algorithm runs at given intervals, individually for each replica. However, data consistency and concurrency control are not considered in DYFRAM. Additionally, DYFRAM doesn't guarantee data availability and system reliability when all sites have negative utility values. In [28], the authors present a horizontal fragmentation technique that is capable of taking a fragmentation decision at the initial stage, and then allocates the fragments among the sites of DDBMS allocation accordingly. The authors in presented a method for modeling the distributed database fragmentation by using UML 2.0 to improve applications performance. This method is based on a probability distribution function where the execution frequency of a transaction is estimated mainly by the most likely time. However, the most likely time is not determined to distinguish the priorities between transactions. Furthermore, no performance evaluations are performed and no significant results are generated from this method. A database tool shown in [30] addresses the problem of designing DDBs in the context of the relational data model. Conceptual design, fragmentation issues, as well as the allocation problem are considered based on other methods in the literature.

However, this tool doesn't consider the local optimization of fragment allocation problem over the distributed network sites. In addition, many design parameters need to be estimated and entered by designers where different results may be generated for the same application case. Our fragmentation approach circumvents the problems associated with the aforementioned studies by introducing

TABLE I: Comparison between Existing Methods in the Literature and the Proposed Approach

Method #	Integrity	Reliability	Communication overhead	Manageability	Data consistency	Security	Performance evaluation
8,11	No, majority doesn't apply clustering and allocation techniques	Poor due to fragments redundancy and Restrictions on the final # of fragments	High traffic to multiple sites, and some delays are tolerated	Difficult due to fragments overlapping and redundancy	Low due to outcomes of the fragmentation	Secured	Intractable complexity are barrier to performance
13,15	No, majority doesn't apply fragmentation and allocation techniques	High because each site is considered in multiple clusters	High due to site duplication in different clusters	Not easy as the # of sites increase and to find the shortest distance algorithms	Inconsistencies are not tolerated	Secured	Poor due to redundant data in multiple sites
19,20,25, 28,29, 30	No, majority doesn't apply fragmentation and clustering techniques	Poor due to fault tolerant risk	Low due to the small # of data allocations	Easy for each site, until there is a need to share data across sites	Presents data consistency for each site, but not across sites	Secured	No Guarantees of high performance as # of sites become intractable
34, 35, 36	No, majority have their own strategy for data allocation, but doesn't apply fragmentation and clustering techniques	High due to the control of cloud services provider	Low as the cost is per use in a unit of time	Easy since it is supported by the cloud service provider	High as it is the responsibility of the cloud service provider	Not secured since it is controlled by cloud services provider	Acceptable level of performance due to more processing required for securing data
IFCA	Yes, full integrated fragmentation, clustering, and allocations	High since data available at most benefit sites	Low due to place data where the least communication cost is hold	Easy because the complete data life cycle can be controlled by the provided techniques	High due to the allocation technique in this approach that keep data consistency	Secured	High as the three-fold techniques guarantees minimum data records allocated to sites with least comm. costs

IV. TELEMEDICINE IFCA COMPUTATION SERVICES AND ESTIMATION MODEL

In the following sections, we present our IFCA and provide mathematical models of its computations' services.

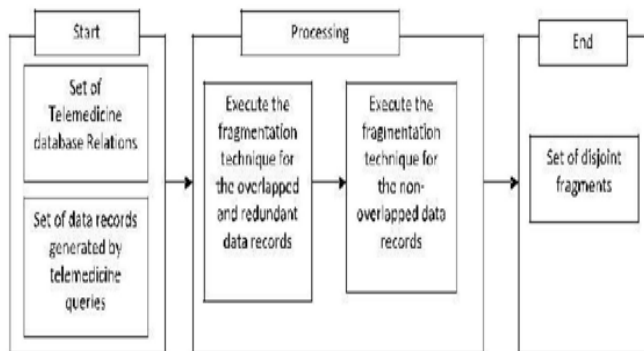


Fig.2. Data fragmentation service architecture.

III. TELEMEDICINE IFCA ASSUMPTIONS AND DEFINITIONS

Incorporating database fragmentation, web database sites' clustering, and data fragments computing services' allocation techniques in one scenario distinguishes our approach from other approaches. The functionality of such approach depends on the settings, assumptions, and definitions that identify the WTDS implementation environment, to guarantee its efficiency and continuity. Below are the description of the IFCA settings, assumptions, and definitions.

A. We Architecture and Communications Assumptions

The telemedicine IFCA approach is designed to support web database provider with computing services that can be implemented over multiple servers, where the data storage, communication and processing transactions are fully controlled, costs of communication are symmetric, and the patients' information privacy and security are met. We propose fully connected sites on a web telemedicine heterogeneous network system with different bandwidths; 128 kbps, 512 kbps, or multiples.

B. Fragmentation and Clustering Assumptions

Telemedicine queries are triggered from web servers as transactions to determine the specific information that should be extracted from the database. Transactions include but not limited to: read, write, update, and delete. To control the process of database fragmentation and to achieve data consistency in the telemedicine database system, IFCA fragmentation service technique partitions each database relation according to the Inclusion-Integration-Disjoint assumptions where the generated fragments must contain all records in the database relations, the original

C. Fragments Allocation Assumptions

The allocation decision value ADV is defined as a logical value (1, 0) that determines the fragment allocation status for a specific cluster. The fragments that achieve allocation decision value of (1) are considered for allocation and replication process. The advantage that can be generated from this assumption is that, more communications costs are saved due to the fact that the fragments' locations are in the same place where it is processed, hence improve the WTDS performance. On the other hand, the fragments that carry out allocation decision value of (0) are considered for

A. Fragmentation Computing Service

To control the process of database fragmentation and maintain data consistency, the fragmentation technique partitions each database relation into data set records that guarantee data inclusion, integration and non-overlapping. In a WTDS, neither complete relation nor attributes are suitable data units for distribution, especially when considering very large data. Therefore, it is appropriate to use data fragments that would be allocated to the WTDS sites. Data fragmentation is based on the data records generated by executing the telemedicine SQL queries on the database relations. The fragmentation process goes through two consecutive internal processes: (i) Overlapped and redundant data records fragmentation and (ii) Non-overlapped data records fragmentation. The fragmentation service generates disjoint fragments that represent the minimum number of data records to be distributed over the websites by the data allocation service. The proposed fragmentation Service architecture is described through Input-Processing-Output phases depicted in Fig. 2. Based on this fragmentation service, the global database is partitioned into disjoint fragments. The overlapped and redundant data records fragmentation process is described in Table 2. In this algorithm, database fragmentation starts with any two random data fragments.

V. CONCLUSION

In this work, we proposed a new approach to promote WTDS performance. Our approach integrates three enhanced computing services' techniques namely, database fragmentation, network sites clustering and fragments allocation. We develop these techniques to solve technical challenges, like distributing data fragments among multiple web servers, handling failures, and making tradeoff between data availability and consistency. We propose an estimation model to compute communications cost which helps in finding cost-effective data allocation solutions. The novelty of our approach lies in the integration of web data-base sites clustering as a new component of the process of WTDS design in order to improve performance and satisfy a certain level of quality in web services. We perform both external and internal evaluation of our integrated approach. In the internal evaluation, we measure the impact of using our techniques on WTDS and web service performance measures like communications cost, response time and throughput. In the external evaluation, we compare the performance of our approach to that of other techniques in the literature. The results show that our integrated approach significantly improves services requirement satisfaction in web systems. This conclusion requires more investigation and experiments. Therefore, as future work we plan to investigate our approach on larger scale networks involving large number of sites over the cloud. We will consider applying different types of clustering and introduce search

based technique to perform more intelligent data redistribution. Finally, we intend to introduce security concerns that need to be addressed over data fragments.

VI. REFERENCES

- [1]J.-C. Hsieh and M.-W. Hsu, "A Cloud Computing Based 12-Lead ECG Telemedicine Service," BMC Medical Informatics and Decision Making, vol. 12, pp. 12-77, 2012.
- [2]A. Tamhanka and S. Ram, "Database Fragmentation and Allocation: An Integrated Methodology and Case Study," IEEE Trans. Systems, Man and Cybernetics, Part A: Systems and Humans, vol. 28, no. 3, pp. 288-305, May 1998.
- [3]L. Borzemski, "Optimal Partitioning of a Distributed Relational Database for Multistage Decision-Making Support systems," Cybernetics and Systems Research, vol. 2, no. 13, pp. 809-814, 1996.
- [4]J. Son and M. Kim, "An Adaptable Vertical Partitioning Method in Distributed Systems," J. Systems and Software, vol. 73, no. 3, pp. 551-561, 2004.
- [5]S. Lim and Y. Ng, "Vertical Fragmentation and Allocation in Distributed Deductive Database Systems," Information Systems, vol. 22, no. 1, pp. 1-24, 1997.

Author's Profile:



Grandhi Lokesh is pursuing M.Tech (CSE) from G.V.R&S college of Engineering & Technology, budampadu, Guntur, AP, India.



Metlapalli Ayyappa Chakravarthi is working as Associate Professor in CSE Department of G.V.R&S college of Engineering & Technology, budampadu, Guntur, AP, India.