IOSR Journal of Computer Engineering (IOSRJCE)
ISSN: 2278-0661 Volume 2, Issue 4 (July-August 2012), PP 05-11
www.iosrjournals.org

A Grid Resource Broker based approach for the integration of Wireless Sensor Networks (WSN) with Grid for Healthcare Applications

Rajesh Eswarawaka Research Scholar of JNTUK, Kakinada
T M Jaya Krishna C.R Engineering College Tirupati, Srinivas Vadali KIET Kakinada
Department of Computer Science and Engineering

Abstract: *Wireless Sensor Networks (WSN) consists of distributed sensor devices. Healthcare application domain is one of the emerging domains in the current world. Nowadays WSN is more popular in healthcare applications and it produces enormous amount of data in a periodic interval. The data should be effectively stored and later it should be processed and analyzed by the doctors to understand the health conditions of the patients. But the main drawback of WSN is it could not able to store large amount of data. Hence there is a need for the scalable environments like Grid to effectively store the data and use later for processing and analyzing the data. To accomplish the objective, in this research paper we have investigated to integrate WSN with Grid environments using the resource broker based approach. The proposed work is integrated with WSN using the Proxy Connector to collect the healthcare data. The collected data is parsed and allocated to the Grid resources using Genetic Algorithm (GA) based scheduling mechanism. The proposed work is aimed to decrease the data transfer time and increase the success rate of data job requests and throughput.*

Keywords: *Grid Computing, Wireless Sensor Networks (WSN), Genetic Algorithm, Healthcare applications, Scheduling.*

I. INTRODUCTION

Nowadays huge rates of elderly people are in need of constant care periodically. For this, they need to move towards the healthcare center on a regular basis, which is not usually accepted by them. In addition, they need to remember the symptoms of their problems. After that, the doctors make a diagnosis and monitor their health. However, some abnormalities only occur in some situations. Hence, it is mandatory to monitor their health abnormalities while they are doing their daily routines at home/office. This can be realized more effectively by adopting Wireless Sensor Networks (WSN) in healthcare applications. The primary advantage of WSN is that it can be deployed to any desired environment, gather and processes the information and delivers it to the interested parties. WSN is an apt solution to address these issues in senior citizens' healthcare. Because the wireless connectivity and sensors size of WSN provides more mobility to the patients. In addition, the wearable sensors of WSN seamlessly collect the data from the patients. Further, WSN increases the quality of services of healthcare services provided to the patient and eliminates the presence of patient at the doctor premises.

In addition, WSN is very different from conventional wireless networks such as Wireless Local Area Network (WLAN), Mobile Ad-hoc NETwork (MANET) or cellular networks. The primary focus of these networks are to deliver high throughput, whereas the WSN is focused towards the energy saving. In general any WSN consists of sensors, base station and a gateway. The sensors can monitor, sample and process the signals. For example, ElectroCardioGram (ECG) sensor can monitor the functioning of heart, ElectroEncephaloGram (EEG) analyzes the functioning of brain and a blood pressure sensor monitors the blood pressure. Then the data is forwarded to the base station. The base station can be a Personal Digital Assistant (PDA), cell phone or a personal computer. The base station is coupled to the server which collects and processes the data. The information in the data is related to the health and environment of the elderly people.

Grid computing is the technology that is mainly aimed to eliminate the scalability and availability issues available in the market for storage and computational resources in an endless manner. It aims to share the computing resources of organizations and individuals across the globe to create a pool of computing or data resources. The Grids are classified into three major types such as Computational Grid, Data Grid and Service Grid. The Computational Grid is responsible for splitting the complex application or job into multiple tasks and run the application using the available computing power. The Data Grid is responsible for storing the large amount of data in the available data storage. In Data Grid a user/application typically submits a job to the Resource Broker that contains the name of the files and the contents the user/application wants to analyze within the stipulated period. But the required files are available in the Grid resources so it is essential to incorporate an effective resource selection

mechanism to select the resources. The Service Grid is responsible for providing the concept of applications as service to the users.

The integration of WSN with Grid using Grid resource broker based approach is shown in Figure 1. Resource Broker is the component that determines how and when to acquire the Grid resources for the applications. There are several brokers are proposed and developed by different commercial companies and academic research groups and it is mainly used in the scientific projects. The integration of WSN with Grid needs an effective resource management for resource discovery and scheduling of data to the Grid resources. The group of WSN nodes is connected to the proxy connector. The proxy connector in turn makes a connector the Grid resource broker. The Grid resource broker is responsible for aggregating the data from the WSN nodes. It first discovers the potential Grid resources that are capable of storing the data. Then it makes the scheduling decisions to allocate and store/retrieve the data into/from the Grid resources. In brief, the contributions of the research work are summarized below:

A. To design and develop the Wireless Sensor Node (WSN) Connector and WSN Data Parser to connect and parse the WSN nodes data.
B. To design and develop the GA based scheduler for allocating the Grid resources to effectively store the data with minimal cost.
C. To design and develop the Data Transfer Manager and Data Replication Manager to transfer and replicate the data in the Grid resources.
D. Integration of proposed (A), (B), (C) to efficiently manage the WSN data store and retrieve requests in the Grid infrastructures.

The rest of the paper is organized as follows: section two highlights related works closely related to our proposed work and some of the background work that supports our proposed work. Section three describes the proposed architecture and the various components in detail. Section four describes the proposed genetic algorithm for data transfer and retrieval requests, section five discuss the implementation details, experimental setup and the inferred results. Finally, six concludes the research work and explores the possibilities of future work.

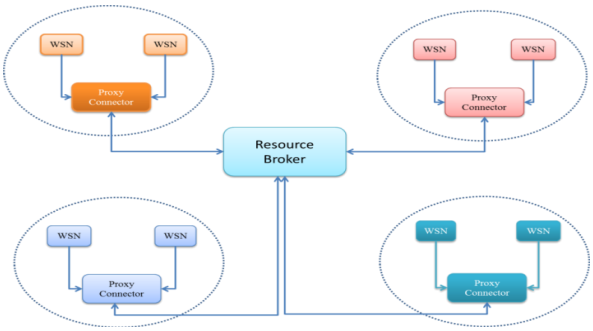

Figure 1: Overview of Resource Broker integration with WSN

II. RELATED WORKS

This section discusses the related works closely related and supports our research work. Recently, research efforts are beginning to study the integration of wireless sensor networks and grid computing. There are several genetic algorithm (GA) based schedulers for computational grid but in our proposed we have make use of the GA based scheduling for data grid. However, most of the GA based schedulers makes the scheduling decisions mainly aimed to maximize the resource utilization, minimize the makespan and flowtime. M. Aggarwal et al. [2] have proposed GA based scheduler for managing the Directed Acyclic Graph (DAG) jobs. Their main objective is to minimize the makespan and idle time of the available computational resources. Y.Gao et al. [2] have proposed two algorithms in the system level and application level to forecast the completion time of jobs in the grid. In the application level scheduling genetic algorithm is applied to minimize the average completion time of jobs through optimal job allocation on each resource. Carretero et al. [3] have proposed a multi-objective genetic algorithm for scheduling mainly aimed to minimize the makespan and flowtime as performance measures. They have represented chromosomes using the direct and indirect encoding and each encoding have different crossover operations. They have used this model to compute the Expected Time to Compute (ETC) on the resources. Their proposed approach is compared with different genetic operators for different number of evolution steps.

The Discovery Net project [4] has built the grid-based framework for developing and deploying knowledge discovery services to analyze the data collected from the sensors for various applications such as life sciences, environmental monitoring and geo-hazard modelling. Zhao et al. [5] propose an efficient monitoring infrastructure for wireless network sensor. Rather than analog to a weather map or air-traffic radar images, the sensor network is used to scan the geographical distribution of network resources and also it provides only an abstract view of the energy resource distribution. Jin et al. [6] has proposed GA based mechanism for energy optimization in wireless sensor networks. They have made use of GA for the formation of a number of pre-defined independent clusters that

helps to reduce the minimum communication distance. Their results have explained that number of cluster-heads is about 10% of the total number of nodes and also they have decreased the communication distance by 80%. Souto et al. [7] is a publish/subscribe architecture for WSNs and they make use of the sensor data if the user has subscribed to the specific sensor data. The messages are aggregated in the cluster nodes for further processing.

III. PROPOSED SYSTEM ARCHITECTURE

The proposed architecture for the integration of Grid resources with Wireless Sensor Network (WSN) is shown in Figure 2. The proposed architecture is associated with three major elements such as Grid resource broker, Grid user and Grid resource provider. The Grid resource broker acts as an entry point to the Grid Execution and Data Management services. It is responsible for scheduling the applications or jobs into the Grid resources, monitoring the jobs and resources. The scheduling is carried out by implementing the suitable scheduling policy that will consider the availability of resources, execution cost, jobs characteristics and etc. Normally the Grid user should be authenticated and authorized user and we consider two types of user. First one is the WSN proxy connector it has been delegated as Grid user and the doctor who is monitoring the patient is another Grid user. The Grid resource provider is the one who owns the Grid resources and it may be computational and storage resources. The resource provider has to make registration with the Grid resource broker for the Grid resources they have owned by them. And also they should get X.509 certificate for the authentication and authorization access to the Grid resources.

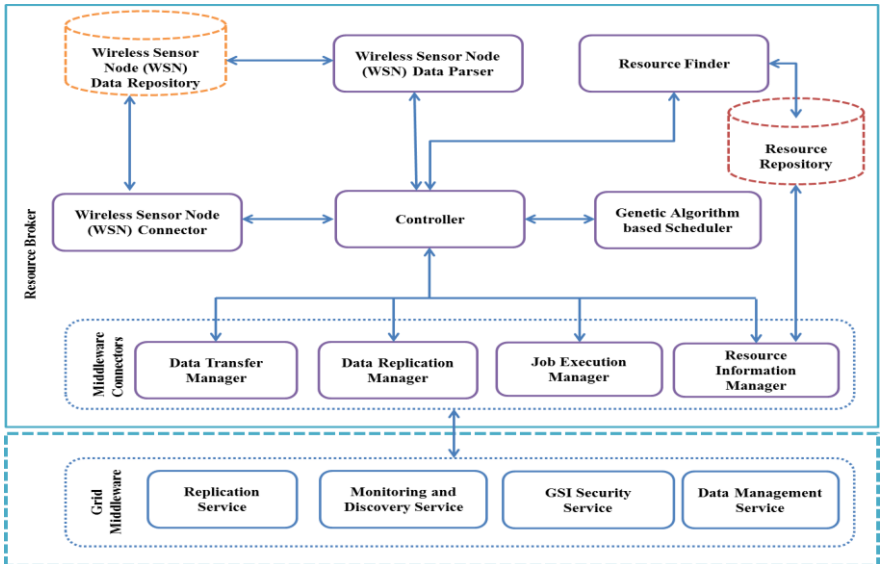

Figure 2: Proposed System Architecture for integration of WSN with Grid Infrastructure

A. Wireless Sensor Node (WSN) Connector

The WSN connector is available in the Grid resource broker. It talks with proxy connector and the data has been retrieved from the proxy connector. It converts the data into XML format for interoperability purpose and stores the data in the WSN data repository temporarily. Once the data has been scheduled and uploaded in the Grid, the temporary data should be removed that resides Grid resource broker.

B. Wireless Sensor Node (WSN) Data Parser

The WSN data parser is responsible for parsing the WSN data that is available in the WSN data repository. The parsed data has been sent to the resource finder to fetch the capable resource to store the data.

C. Controller

The Controller module is responsible for the invoking the following components such as

- Resource Finder to finds out the suitable resources
- GA based scheduler to select the best resource to store the data
- Data Transfer Manager to transfer the data to the Grid resources
- Data Replica Manager to replicate the data in the Grid resources

D. Resource Finder

The resource finder fetches the data requirements and matches the requirements with available Grid resource information in the resource repository. It fetches the set of resources that are capable of solving the tasks.

E. Genetic Algorithm (GA) based Scheduler

Scheduling is a key component to schedule the user application requests to appropriate services or resources. Our scheduling mechanism is based on genetic algorithm that has the capability of allocating the data requests in a near optimal manner. The main idea of the scheduling is to allocate the data job requests to the available resources in an efficient manner. To achieve this, GA based approach finds out the best match from the list of available jobs to the list of available resources. The proposed GA based scheduling mechanism has been explained detail in the next section.

F. Data Transfer Manager

The data transfer manager is responsible for transferring the parsed WSN data resides in the Grid resource broker to the Grid resources. It makes use of the reliable file transfer service for transferring the files to the appropriate resource provider.

G. Data Replication Manager

The data replication manager is responsible for replicating the data contents into multiple Grid resources. It replicates the data by considering the load of the data in the Grid resources.

H. Job Execution Manager

The job execution manager is used for managing the computational nodes in the Grid It analyses the job flow and select the computational resources to process the job. It makes use of the Globus Resource Allocation Manager to select and allocate the computational resources.

I. Resource Information Manager

The resource information manager is used to aggregate the various Grid resource information such as number of nodes, processor architecture, operating system, ram storage, hard disk capacity, processor speed, processor utilization, bandwidth and latency.

J. Reliable File Transfer (RFT) Service

The RFT service is mainly aimed to achieve the reliability to transfer the data in a reliable manner using the by check pointing the data in the postgresql database. It has provided with source and destination URL's to transfer the data from the Grid resource broker site to the Grid resource site.

K. Replica Location Service (RLS)

The replica location service (RLS) maintains and provides access to mapping information from logical names for data items to target names. These target names represent the physical locations of data items, or an entry in the RLS may map to another level of logical naming for the data item. The RLS is intended to be one of a set of services for providing data replication management in Grid.

The resource brokering algorithm used to perform various operations for the integration of WSN with Grid resources is explained in Algorithm 1.

Algorithm 1 Resource Brokering Algorithm

Input: Set of data store requests from the WSN nodes monitoring health care data and data retrieve requests from the health care application monitored by the doctors.
Output: Store or Retrieve the data from the Grid resources.
Pre-requisite: Start the resource broker to retrieve the Grid resource information by interfacing with the Grid middleware.
Step 1: Get and parse the data store or retrieve requests from the WSN nodes and doctors then store the requests as objects.
Step 2: For every application request whether it may be store or retrieval operation perform the matchmaking with the available Grid resources
Step 3: Find out the matched resources in terms of storage memory for every application request and sort the resources based on the rank.
Step 4: For every job request with matched resource list apply the GA based scheduling and select the best resource based on the fitness value to store or retrieve the data.
Step 5: Invoke the Data Transfer Manager to transfer the data to the Grid resources or retrieve the data from the Grid resources.
Step 6: End For

IV. GENETIC ALGORITHM (GA) BASED SCHEDULING

We have used Genetic Algorithms (GAs) for scheduling the data to the Grid resources. GA is a kind of Meta-heuristic approach which is being used to tackle the problem of grid scheduling. Genetic Algorithm (GA) [ ] is stochastic in nature, domain-independent search, and optimization techniques and it has been developed in 1970s. It has borrowed the concept of nature selection that is based on the principle of stronger survives in the competition and the weaker will decease so that the genes available contained within the chromosomes of dominant individuals will spread within the next generation. In our proposed GA, we have concentrated upon the performance of the Grid Scheduling by taking a few parameters into consideration i.e. data transfer time, successful data job requests and throughput. The main objective of our proposed work is to minimize the data transfer time and increase the successful data job requests and throughout.

4.1. The proposed Genetic Algorithm (GA)

The proposed algorithm retrieves the job requests from multiple WSN nodes and generates an initial population. Then each chromosome in the generated initial population has been assigned to randomly selected Grid resources and generates the job and resources pair. If the number of jobs in the queue has been less than the chromosome size then the first 'n' chromosomes has been used in solution encoding and the chromosome size has been set to n. Then the proposed algorithm determines and allocates the best schedule by applying the genetic operators such as crossover and mutation. Once the best schedule has been obtained the jobs has been dispatched from the job queue and dispatched jobs has been removed from the job queue. The modified GA algorithm performs the following operations such as chromosome representation, genetic operations of mutation, crossover, and fitness value and it is described below.

Let $T_{TakenJobs}$ be the job list taken in the scheduling interval S. Let R be the number of Grid resources available. The resource optimization algorithm for each T_{Jobs} matches with the available grid resources R. If more than one grid resource R has been suitable for submitting a particular job the matched resources has been stored in $J_{MatchList}$. Let chromelength represents the length of chromosome, pop size is the size of the population, T_{Jobs} represents the total number of jobs in the chromosome and it has been initially set the value of 0, $T_{Jobs[i]}$ represents the i_{th} job in the job queue. The proposed algorithm to perform GA based scheduling is given in Algorithm 2.

Algorithm 2 Proposed GA based Scheduling Algorithm

```
Genetic Algorithm ( )
{
Begin
If (T_TakenJobs. Size ( ) >0)
{
For (int II =0; II <pop size; II++)
For (int JJ=0; JJ < T_OrderedJobs. Size (); JJ++)
{
First generate a random number of pop sizes from (1 to R)
Second randomly select a position of q (1 to T_Jobs + T_TakenJobs)
Third insert the job (T_Job [i], pop size) at position q of chromosome
}
}
If (TJobs .Size ()> 0)
{
Evaluate the fitness member function
For (int i; i < TJobs.Size (); i++)
{
Perform crossover operation
Perform mutation operation
}
Select the best chromosome in the population as a solution
}
Else
{
Return 0;
}
End
}
```

V. IMPLEMENTATION DETAILS EXPERIMENTAL SETUP AND RESULTS

5.1 Implementation Details

The proposed architecture is implemented using SOAP based web services to avoid the interoperability and portability issues it is deployed in Sun Glassfish Server 3.0.1, MySQL 5.0.23 is used as the database. The WSN connector service fetches the data from the proxy connector in the group of WSN nodes. The fetched data is parsed by the WSN data parser module and parsed data has been updated in the WSN data registry as a XML format. The controller service in a periodic interval invokes the Resource Finder service to match the user application requests with the available Grid resources for storage space, bandwidth and latency. The controller service invokes the GA based scheduler to select the resources from the matched resource list. The GA based scheduler service selects the best resource based on the objective of minimal transfer cost. Then the controller invokes the Data transfer service to transfer the data to the Grid resources. If the data has to be replicated the controller invokes the Data replication manager service to replicate the data in the Grid resources. The GA based scheduler and Resource Finder makes use of the discovered Grid resources information by the Resource Monitoring service. The Resource Monitoring updates the Grid resource's information in XML format.

5.2 Experimental Setup

The following experimental setup is made in our campus for testing the proposed work and it is shown in Figure 4. It consists of five clusters namely jntuclusterwsn1.care.mit.in, jntuclusterwsn2.care.mit.in, jntuclusterwsn3.care.mit.in, jntuclusterwsn4.care.mit.in and jntuclusterwsn5.care.mit.in. Each cluster has one head node and 4 compute nodes respectively. For testing the proposed work we have installed the jntuclusterwsn1.care.mit.in and jntuclusterwsn2.care.mit.in in High-Performance Computing (HPC) laboratory and jntuclusterwsn3.care.mit.in, jntuclusterwsn4.care.mit.in and jntuclusterwsn5.care.mit.in in WSN laboratory. The Globus Toolkit 4.0.3 is used as Grid middleware, Ganglia version of 3.2.x is used as resource monitoring tool and NWS version of 2.13 is used as network monitoring tool. The proposed mechanism is tested by creating 20 WSN nodes in the HPC laboratory that will generate the WSN data in a periodic interval of 10 minutes. We have configured four WSN proxy connector and every five WSN nodes have been connected to one proxy connector and in turn it is connected to the Grid resource broker.

5.3 Experimental Results

The experiment been carried out for testing the GA based scheduling with Grid resource broker. The input for the algorithm is job requests that consists data size of the file, matched job resource list for each data job requests. The performance of the proposed algorithm has been compared with the conventional scheduling approach for the data transfer time. The jobs have been tested by generating in a random fashion 20 to 100 jobs with the interval of 2 to 3 minutes. The comparison of data transfer time based on the proposed approach with conventional approach is shown in Figure 3.

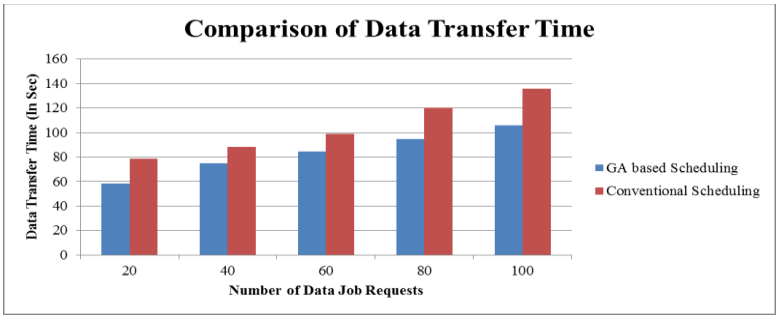

Figure 3. Comparison of Data Transfer Time

The performance of the proposed algorithm has been compared with the conventional scheduling approach for comparing the successful data requests by the proposed approach. The same set of jobs has been tested by generating in a random fashion 10 to 100 jobs with the interval of 2 to 3 minutes. The comparison of data transfer time based on the proposed approach with conventional approach is shown in Figure 4.

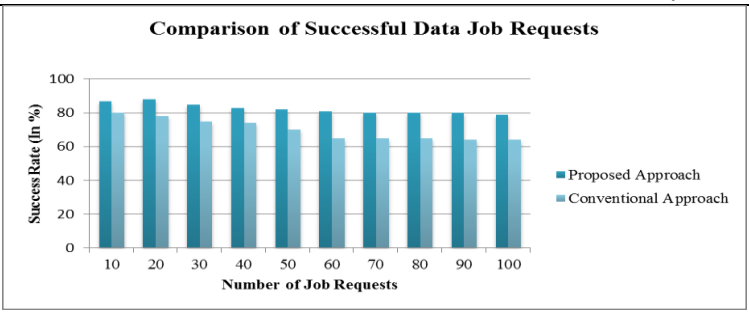

Figure 4. Comparison of Successful Data Job Requests

The resource broker efficiently discovers the suitable Grid resources based on the user application requirements and store the data job requests in the Grid resources. The same set of applications has been submitted, and the application throughput has been compared for various test cases. The proposed mechanism increases the throughput of applications submitted to resource broker, and it is shown in Figure 5.

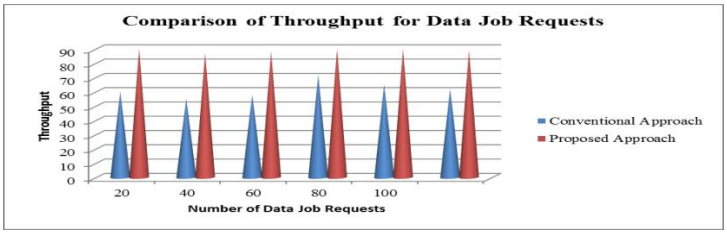

Figure 5. Comparison of throughput of applications

VI. CONCLUSION AND FUTURE WORK

The Grid is a complex environment with a large number of computational resources from different organizations collaborating to solve large scientific applications. In recent years, the grid concepts have been evolved to not only satisfy specific domain but applications from diverse domains. Hence, the scope of grid increases resulting in participation of more number of resources and organizations, and thus management of grid resources becomes complex. This paper mainly aimed to proposed to develop a Grid resource broker based approach for the integration of WSN with Grid resources. The proposed approach efficiently stores the WSN data in the Grid resources in a scalable and secured manner. The integration of GA based scheduling with Grid resource broker effectively select the Grid resources that will minize the makespan and flowtime of the requests submitted to the Grid resource broker. The proposed system enhances the performance measures such as data transfer time, successful data job requests and throughput. As a future work, WSN nodes for healthcare application sometimes generate critical data this type of data has to give more priority during the selection process.

References

[1] I. Foster et al., The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration, Grid Forum white paper, 2003

[2] J. Carretero, F. Xhafa, and A. Abraham, "Genetic algorithm based schedulers for grid computing systems," International Journal of Innovative Computing, Information and Control (ICIC), vol. 3, pp 1349-4198, 2007.

[3] M. Aggarwal, R. D. Kent and A. Ngom, "Genetic Algorithm Based Scheduler for Computational Grids," in Proc. of 19th IEEE International Symposium on High Performance Computing Systems and Applications, 2005.

[4] Y. Gao, H. Rong, and J. Z. Huang, "Adaptive grid job scheduling with genetic algorithms," Future Generation Computer Systems, vol. 21, pp.151-161, 2005.

[5] Discovery Net Project, http://www.discovery-on-the.net.

[6] S. Jin, M. Zhou, and A. S. Wu, "Sensor network optimization using a genetic algorithm," in Proceedings of the 7th World Multiconference on Systemics, Cybernetics and Informatics, 2003.

[7] Souto E., et al, Mires: A publish/subscribe middleware for sensor networks. In ACM Personal and Ubiquitous Computing, 10(1): 37-44.

[8] E. Jovanov, "Patient Monitoring using Personal Area Networks of Wireless Intelligent Sensors," Biomedical Sciences Instrumentation, vol. 37, 2001, pp. 373–378.

[9] Y. Yao and J. Gehrke, "The Cougar Approach to In-Network Query Processing in Sensor Networks." ACM SIGMOD Record, vol. 31, no. 3, ACM Press, 2002, pp. 9–18.

[10] M. Gaynor, S. L. Moulton, M. Welsh, E. LaCombe, A. Rowan, and J. Wynne. Integrating wireless sensor networks with the Grid. IEEE Internet Computing, 8(4):32–39, August 2004.

[11] B. Krishnamachari, D. Estrin, and S. Wicker. The impact of data aggregation in wireless sensor networks. In Proceedings of thr IEEE 22nd International Conference on Distributed Computing Systems Workshop, pages 575–578, July 2002.

[12] Lim H. B., Teo Y.M., Mukherjee P., Lam V.T. et al. 2005. Sensor Grid: Integration of Wireless Sensor Networks and the Grid, In Proc. of the IEEE Conf. on Local Computer Networks.