CHAPTER 2

MOBILE GRID

A mobile grid extends the capabilities of a grid to mobile users. The participating nodes in such a grid are static as well as dynamic. This chapter discusses the motivation for a mobile grid, its dimensions and related topics such as wireless mobile networks, smart devices, and context-aware applications. This chapter also includes a potential case study related to healthcare that could benefit from resource discovery in mobile grids.

2.1 MOTIVATION

A key requirement of a grid computing infrastructure is to support its users their diverse requirements of computing devices. Primarily grid is required to support different needs of different users at different times for different applications. Mobility has seminally influenced the evolution and design of distributed systems (Satyanarayanan 1996) and grid is not an exception. Mobile computing is very mature in technology and the adaptation of mobile devices in computing is enormous.

Wireless devices have become an indispensable part of life. The reason for such adaptation of mobile devices is manifold. From the user’s point of view, with the size of the device shrunk, these devices can be carried out with ease, thus promoting mobility. Some of these devices have computing capabilities on par with their static counterparts. From the application point of view, mobile devices have special types of applications including mobile business transactions, location-based services and context-aware applications. Thus from the user and application points
of view, it is favorable to integrate mobile devices in a grid computing infrastructure.

From the *middleware point of view*, the mobile devices pose serious challenges (Badrinath 1993, Satyanarayanan 1996) like, scalability, heterogeneity, connectivity, security, and fault tolerance in the presence of node mobility and node dynamism. Despite these issues mobile devices are integrated into a grid to take advantage of the combined use of these two technologies. Such integration offers the user computing capability and data access anytime, anywhere.

Devices such as Laptops with wireless technology, PDAs enabled with Bluetooth or infrared interfaces, mobile phones etc. are few examples of mobile devices. A majority of these devices have reduced performance, limited storage, unreliable bandwidth communication. Nevertheless, these devices include value-additions such as cameras, microphones, GPS receivers, satellite receiver/transmitter, RFID tags, special-purpose sensors, etc. (Manvi 2010). Special applications can be realized cost-effectively due to resource sharing by integrating these devices in the grid environment. Further, these devices can overcome their limitations by deriving benefit from joining a grid where they have the opportunity of using a number of resources connected by a network.

### 2.2 WIRELESS MOBILE NETWORKS

In comparison with traditional (wired) grids, mobile grids have the advantage of exploiting the concepts of mobility and ubiquity in term of being available any-time, anywhere. Making use of wireless connection, the mobile grid environment is available any-time and any-where. Wireless connectivity is obtained by different types of wireless networks classified based on either topology or range and coverage as shown in figure 2.1.
Wireless technologies include Wireless LAN (IEEE 802.11 standard), cellular technologies, paging systems, satellite systems, Bluetooth, ZigBee, etc (Andrea 2005). Bluetooth radios provide short range connections between wireless devices along with basic networking capabilities. Its normal range of operation is 10m (1 mW transmit power) to 100m (100mW). The ZigBee radio is designed for low cost and less power consumption compared to that of Bluetooth.

Paging systems broadcast short messages simultaneously from many tall Base Stations or satellites transmitting at very high power. The advantage of low cost the pagers had was challenged by the competitive cost of cell phones. Satellite systems are bigger, consume more power and obviously expensive although they offer ubiquitous worldwide coverage. So they are deployed for sectors like broadcast entertainment. Wireless LANs (WLAN) provide high-speed data rates, but the coverage area is limited. Despite this issue, WLAN is extremely popular because of applications like Internet access, emails requiring lesser bandwidths and web-browsing, etc.
Cellular networks are adopted extensively with global deployment of next-generation cellular technologies like General Packet Radio Service (GPRS), Universal Mobile Telecommunications System (UMTS), Code Division Multiple Access (CDMA), WiMax, and Long Term Evolution Networks (LTE). Grid users can benefit from advances in cellular networks which give users seamless wireless coverage, broadband access in hot-spot areas and ubiquitous coverage everywhere. This thesis therefore assumes mobile grids built on cellular networks.

2.3 CELLULAR NETWORKS

A cellular network consists of mobile units linked via a radio network to the Public Switched Telephone Network (PSTN). The service coverage area is divided into multiple cellular regions each populated by Mobile Nodes (MN) and governed by a Base Station (BS). BS is fixed and is connected to Mobile Switching Centre (MSC) which in turn is connected to PSTN. Thus MSC controls a cluster of BS. The wireless link between the MN and BS enables MN to communicate with wired phones in PSTN. Figure 2.2 depicts how a MN interacts in a cellular network.

![Cellular Network Diagram](image)

**Fig. 2.2 MN linked to PSTN via BS and MSC.**

When the MN is active it identifies and registers with an appropriate BS based on its location. Its cell position is stored at the corresponding MSC. The BS monitors the signal quality and reports it to the MSC which in turn decides about
routing of the call. When the MN moves from one cell to another, the BS senses deterioration in signal quality and informs MSC about it. MSC in turn, switches the control of the call to the BS of the new cell where the MN is located. This procedure is called handoff. A number of approaches to handoff (Zeng 2002) are available and each one has a different impact on the on-going computation and communication. The grid middleware considers the impact as a challenge especially when required to satisfy the QoS expectations of the users.

2.4 SMART DEVICES

Mobile computing involves software, hardware and mobile communication (Masoud 2012). Wireless grid applications are grouped into three categories (McKnight 2004) namely, (1) applications collecting data from nomadic devices, (2) location-related applications, and (3) applications using wireless network connected devices. These applications can be implemented using wireless sensor networks and wireless ad hoc networks.

Cellular networks have been widely used to provide connectivity for smart devices employed in these types of networks. Smart devices can be personalized and configured to a specific owner. Smart mobile device access is characterized by open service discovery and intermittent resource access. Smart devices can be designed to support a variety of applications especially related to ubiquitous computing. Such devices when integrated with sensors can deal with a number of critical applications.

A mobile grid built on a cellular network can support such sensor-based applications. Mobile grid being a revolutionary paradigm shift, enables applications like wireless sensor networks, automated highways and factories, e-learning, smart homes and appliances, remote telemedicine, disaster management, automobile
sector, etc. Of various applications that can be implemented in a mobile grid, this thesis presents a context-aware application in the field of medicine.

2.5 CONTEXT-AWARE APPLICATIONS

With the recent advances in wireless communication and portable mobile devices, users are able to access information services through a shared network infrastructure like Internet, regardless of their physical location. The ongoing research in wireless sensor networks has made available a variety of sensors. These sensors are characterized by their sensing, computation and communication capabilities.

A mobile sensor grid is an application-driven system architecture that can be used in a number of applications including the medical field. Special sensor devices play a lead role in collecting huge streams of physiological data representing contextual information. The networked mobile devices can be employed to carry this context information from the sensors to the application logic. This integration facilitates hosting a number of applications including the ever-demanding area of health and medical services.

The wireless sensors planted inside the patient’s body or applied on the surface of the body can collect the necessary real-time data for decision-making. With the inclusion of sensors the mobile grid becomes a mobile sensor grid. The name is justified, as the sensors are also mobile because of the moving nature of the patient. In such a mobile sensor grid, the moving resources possess limited capabilities compared to their static counterparts.

The sensors are expected to pass the relevant data to the application logic via a mobile device only. In doing so, there are several constraints intrinsic to
mobile devices (Satyanarayanan 1996) of which the varying connectivity factor is an inhibiting one for this sample context-aware application.

2.5.1 Health Application

Of various medical care services that can be offered with a mobile sensor grid, this work examines the case of providing continuing care for liver transplant patients. Liver transplant is the treatment of choice for selected patients with end-stage liver disease.

The motivations for taking up this medical application are manifold. Firstly, liver transplant owing to various reasons lacks specialists especially in the developing countries. This requires that the patient after the transplant has difficulty in consulting the physician who is geographically separated. Secondly, Liver Transplantation despite receiving overwhelming success faces critical problems related to its use and continued success. For example, the patient receiving a liver transplant needs to be under lifelong immune-suppression (as of today’s medical research) which itself is associated with its own set of risks and complications (Paul 2001).

It is imperative that such patients need to be constantly monitored and given varying dosage of medicines depending on several tests conducted periodically with blood samples taken from the patient. Follow-up care after liver transplant is crucial in order to prevent infection or rejection of the new organ.

A mobile grid can be employed to help the liver transplant patients, in particular the children who need to undergo immense pain and ordeal of giving the blood samples regularly for receiving appropriate dosage of medicines for immune-suppression. The patient is more likely to get infected after the transplant.
Accordingly, some special sensor devices can be planted inside the body of the patient and be made to monitor and collect data from the blood stream continuously.

Any variation in the dosage therefore can be estimated on receiving data from the sensors by the grid. Through this, there is a possibility that the patient will receive a lesser dosage as required instead of consuming the same over-dosage that was prescribed based on the previous result of periodic monitoring of data. This will in turn help in bringing down the side-effects caused by such medication as unfortunately the therapy is associated with undesired effects (Paul 2007).

Even though it is possible that implanted blood sensors could be developed to monitor the health of liver transplanted patients, non-invasive sensors for such patients are available. FibroScan (FS) is a hand-held instrument. Instead of detecting the presence of immune rejection in blood, it monitors the transplanted liver for any deterioration. It basically measures the Liver Stiffness (LS). Quoting from Sebastin Mueller and Laurent Sandrin (Sebastian 2010), we have:

…..a horizontal position with normal breathing yields the lowest and most reproducible LS values.

…..the major problem is that serum markers reflect the profibrogenic or profibrolytic activity, but do not yield any information about the net deposition of matrix in the liver which are not necessarily correlated to each other.

…..In a post-transplant study of patients infected with HCV, median LS at months 6, 9, and 12 were significantly higher in rapid fibrosers as compared to slow fibrosers. The slope of LS progression in rapid fibrosers was significantly greater than in slow fibrosers, suggesting two different speeds of liver fibrosis progression.
Multivariate analysis identified donor age, bilirubin level, and LS as independent predictors of fibrosis progression and portal hypertension in the estimation group. Another study suggested that TE is a reliable tool to assess liver fibrosis in patients with recurrent HCV after living donor liver transplantation.

FS applies TE (Transient Elastography) to measure LS. From PadGadget (PadGadget 2011) we find that FibroScan Library is an iPad application released on 28th of July 2011. Direct integration of FibroScan to a mobile device is not far off.

2.5.2 Context-Awareness

The health application discussed above is certainly different from traditionally designed distributed applications. It is the ability to implicitly sense and automatically cater to user’s needs that distinguishes context-aware applications from other applications. Pervasive computing has paved way for building such intelligent applications. Applications become smart when they are context-aware. Such application detects and reacts to environment variables.

Context is defined as any information that can be used to characterize the situation of an entity. An entity is a person, place or object that is considered relevant to the integration between a user and an application. In other words, context represents a set of situations and actions including the user and the application themselves. Such situations force to change over time, describing humans’ behaviors, application and environmental states, whenever specific actions are applied (Christos 2007). Context thus refers to both static as well as dynamic data.

The behavior of the application is determined by the information obtained from the sensors. It is analogous to a mobile phone changing proactively its time on entering a new time zone. Thus the application in future should enable the patient consume the right dosage of medicine without concerning itself how the
data was acquired, stored and sent for processing. Context in this application refers to the patient’s biological data and also the data acquisition environment comprising sensors and the mobile devices that act as data collectors.

The sensors collect the data periodically or on command by the patient and transmit the same to the nodes that are a part of the grid. In a mobile grid nodes can be either static or mobile. If it is a mobile node, there is a flexibility of acquiring the data at any place in the grid. FibroScan is portable and hence the place and time a LS would be sent for analysis is unknown. At the same time the situation is more demanding because of the overhead involved in it.

The sensor constantly tracks the mobile node for passing the data or the mobile device constantly queries the sensor. Sending the data or querying for the data could be at random times also. All of these account for communication overheads. In the next section an appropriate framework for hosting this application is discussed.

2.5.3 Design rationale

2.5.3.1 Assumptions

It is assumed that the mobile sensor grid designed has a high availability of resources round-the-clock including persistent network connectivity, as this being a crucial application. That is, to meet the demands of the application in real-time, grid is equipped with a vast collection of required resources. Also, as the application is very sensitive and demanding it is also assumed that the wireless sensor data captured is passed onto the grid with less latency and on a high bandwidth network. The sensor devices in future are assumed to be robust in order to support the cause as the patient is expected to be even unaware of their presence, which makes them continue with normal activities and changing environment.
Security, privacy and admission control issues that are key concepts of any distributed application are not addressed here. Also the type and nature of sensors is not in the scope of this work. However, it is assumed that such sophisticated sensors capable of capturing data related to the enzymes levels in the blood are expected to be available in future.

2.5.3.2 Goals

There are six major goals in the mobile grid for this application, which are discussed as follows.

1. *Activate necessary relief transparently to the patient:*

   The existing system of medical care for liver transplant patients requires that the patient give a blood sample periodically, say every day during the first two months after the transplantation, every week between two to three months and every month later on. The blood sample is analysed with a series of tests in a specialised laboratory and the report is sent to the physician who in turn prescribes the dosage for a specified interval of time. The proposed system does not even remind the patient of this routine and thus relieves him physically as well as psychologically.

2. *Continuous monitoring of blood sample:*

   The sensors planted inside the body of the patient constantly collect data and thus alert any abrupt change as apparent from the blood stream inside the body.

3. *Requires no physician’s direct involvement:*

   As the expertise available in liver transplantation and post-transplantation care is limited, quite often the communication between the patient and the physician becomes infeasible especially when the patient is from a rural front. The sensors directly communicate with the grid and so the system itself will take care of prescribing appropriate dosage without the actual involvement of the doctor in most occasions. However, if any trace of complication or potential
problem, wherever it is warranted the doctor may be consulted or alerted by the grid.

4. *Integrate data from heterogeneous fixed and mobile devices and sensors:*
   The grid should be capable of receiving the data, store, integrate and interpret the same.

5. *Dynamic estimate and dosage:*
   The application logic based on some threshold values available in it, analyses the data it receives. The suitable dosage is prescribed thereafter.

6. *Communication of dosage variation:*
   The patient is informed of any change in the dosage which he is advised to take up subsequently via an appropriate medium, say a display device or a short message service whichever the patient is in possession of.

### 2.5.4 System Design

A context-aware application uses context to provide services to the users. The future health maintenance application involves sensors planted in the patient’s body that capture the patient’s enzyme level data from the blood. The sensors are programmed so that they constantly keep collecting this context data. This data is raw and may also contain data that is not relevant to the application. From the sensors the data is carried to the available mobile device in the grid. Adopting any of the following two strategies can do this. Firstly, the mobile device is selected based on its proximity to the patient. In other words, in order to reduce the communication overhead, the sensors need to trace the nearest neighbor to whom the data can be passed on to.

In a majority of the cases, it can be a mobile device possessed by the patient himself. But it is not expected by the patient to carry his mobile device
always with him. The patient is expected to carry out his normal activities be it travelling or of course even swimming. Alternatively, the sensors can multicast this data to a group of devices in the grid. The device, which first receives the data, sends it for processing. This strategy promotes reliability of this time-sensitive application. Therefore the trade-off between communication overhead and reliability is to be determined to choose one of these strategies. The sensor devices act as data collectors and the mobile nodes act as data carriers for processing.

In line with the vision of aiding the transplant patients a mobile sensor grid can be built as an integration of data, computational and service grids. The grid will constitute both static as well as mobile devices as the latter can also act as service providers and service consumers. The data grid is a grid system that deals with controlled sharing and management of distributed data. For this application, the data that is required for deciding the dosage of medicines will be made available and maintained by the data grid. It is to this grid that the various sensors planted in the body of the patient pass the data to.

The sensors observe the relevant data continuously and this raw data is passed to the available mobile device in the grid. This mobile device, which is a component of the mobile sensor grid in turn, communicates through its Base Station to a dedicated grid node, which is equipped with special blood analysis equipment. This equipment is responsible for analyzing the data. The analysis results are sent to the data grid. The blood analysis results received by the data server need to be structured for our application. The data server following the Open Grid Services Architecture (OGSA) (Foster 2003) does this. OGSA defines basic architectural structure and mechanisms that can be applied to share, interpret, and analyse data collections. The data server not only receives the new data and converts it into a structured form, but also maintains a profile of all patients.
The application server consults this database and takes the specific patient’s data. For every patient attached to the grid, the patient’s history taken from the database is taken as a reference data against which the results obtained after the blood analysis is compared. Accordingly the next dosage is computed. The new dosage is sent to the patient via Short Message Service over the mobile device or any other display device available with the patient. This application has only moderate volume of data compared to some scientific applications and therefore quick processing is ensured.

The physician’s computer or the mobile device is also a part of the grid. Whenever the application finds a remarkable variation in the data it calls for the intervention by the concerned physician. The physician in turn can respond with some other prescription or any other medical advice based on the severity of the situation.

2.5.5 Overview of Challenges

Context-awareness is an active area, which has attracted much research. It becomes more significant especially when resource-limited mobile devices are integrated in the design of health maintenance of liver transplant patients. Context here refers to information pertaining to the patient, sensors, mobile device and the environment. Such context information provides valuable information to the application logic for decision-making. Nevertheless, the following challenges arise in order to make the decision-making process foolproof.

1. Context presentation: The context data collected via the sensors is only raw data. This data may have some unnecessary data that is not relevant for our application. A mechanism, which separates the relevant data and presents in a predefined format is essential.
2. *Speed of communication of context data:* The communication delay incurred in the transfer of sensor data for further processing may affect the performance of the entire system. That is, the system in turn may not be able to compute the right dosage in time. Also, these types of contextual data can change in a short period of time, and the data itself may be incorrect or irrelevant.

3. *Relevance of data:* The sensors employed for capturing the data required for dosage estimation suffer from inherent limitations. For example, recharging of wireless sensors is same as that of mobile devices. Further, the data collected from the sensors travels through a series of devices before it is actually used by the application. Therefore, any delay or loss or disruption of the data questions its validity and may cause a severe damage.

4. *Mobility:* Mobility refers to both the movement of the patient as well as the device. Mobile devices primarily built on a cellular network may enter or leave a cell any moment of time. This type of a temporary disconnection needs to be taken care.

5. *Information to the patient:* The new dosage information is delivered to the patient through a mobile device. Context plays a significant role when we attempt to use such a device. The inherent limitations of these devices need to be given a serious thought.

Having discussed a motivating example for mobile grid, in sequel a basic framework for mobile grid to realize these types of applications is discussed.

**2.6 MOBILE GRID ARCHITECTURE**

Mobile devices in a grid may serve as resource providers, consumers, or both. In order to accommodate mobile devices in a grid, the concept of VOs (Foster
1999) needs to be redefined as in (Martin 2006). Mobile dynamic Virtual Organizations (MDVOs) are an extension to VOs whose members are able to change locations while providing or consuming services even after temporary loss of reachability. Further, not all members have to be mobile in order to denote a VO as an MDVO. Fig.2.3 shows a grid with a single VO and two MDVOs consisting of static nodes (SN) and mobile nodes (MN). It can be observed from the figure that to qualify as MDVO, a VO may have exclusively mobile nodes or a collection of both static and mobile nodes.

Fig. 2.3 A Grid with VO and MDVOs

A layered approach to mobile grid organizational model is defined in (Waldburger 2006). It distinguishes VO as a base VO and operational VO. The base VO has a layered approach comprising pools of potential resources, services, applications, and providers. The basic resource pool constitutes typical resources for
computational and data grids. These resources are encapsulated by a service layer on top. Applications consist of one or more grid services. The provider pool has providers that offer grid services of these applications. An operational VO represents such a base VO and the user with the device along with higher level resources.

2.7 CONSTRAINTS AND ISSUES

The two key characteristics of mobile grids that deserve special attention are multi-domain high-level resource coordination and mobility-triggered dynamics of the organizational composition and business flows. These characteristics and other limitations of mobile devices pose a number of unique challenges that must be overcome when building grid applications using these devices. Examples of other limitations that hinder the integration of mobile devices into grid are relatively poor resources (in terms of computational speed, memory), battery constraints, unreliable connectivity, weak security, etc. (Sang 2003).

Various issues specific to a mobile wireless grid are examined in literature. The following issues or challenges are to be addressed to take the full benefit of a mobile grid.

- Mobile devices are resource-poor relative to static elements.
- Mobile devices are dependent on their battery power for their functioning.
- Mobile connectivity is highly variable in performance and reliability.
- Mobility affects the task execution time and thus is a critical factor affecting the mobile grid performance. If the devices are highly mobile, they affect the computationally-intensive, long-running tasks. If the node leaves the cell while partial tasks are still in progress, it has an adverse impact on the computation.
- Mobility is closely associated with the issue of Quality of Service (QoS). The goal of a mobile grid is to provide the same QoS to mobile users as experienced by wired users. Soft handoffs (Qing 2002) help in improving wireless QoS. The inherent unstable nature of the wireless network causes instability in the network. Frequent disconnections prevalent in mobile grid may result in poor QoS. If such disconnections are anticipated and planned for, mobile devices can expect better levels of QoS.
- The nature of the wireless medium and mobility of devices cause network instability. Techniques to balance network loads are significant towards such temporary, but bursty imbalances.
- Interference can be caused in the wireless channels of communication due to atmospheric disturbances like electromagnetic storms and environmental factors.
- Mobile grid being a wide-area distributed system, it needs to be structured around an asynchronous model only. For example, Remote Procedure Call (RPC) in a synchronous system may have more communication latency than the processing speed. Hence it is better to build a mobile grid as an asynchronous system.
- At any time the grid may revoke resources that it has temporarily delegated to an application when there are higher priority jobs or when the owner of the resources returns to computation. Although this is common in any distributed system which harnesses the idle resources, this issue is severe in a mobile grid because of high mobility. The revocation strategy needs to consider priority of applications. Cost of revoking the same resource may be different for different applications. There may be dependencies between processes that need to be considered during revocation. Revocation strategy has to consider all these issues and failing to do so may cause hazards like deadlock and starvation.
The issue of data loss can be critical when more than one mobile device participates in task-sharing. Since an intensive task is divided into small partial tasks for distribution across the grid, latency in handoffs is a key issue. In a resource-sharing environment such as the wireless grid, loss of data can result in potentially erroneous outcomes.

This chapter has given a detailed view of a mobile grid, motivational application, architecture and various challenges that need to be addressed. In the next chapter building a mobile grid as a P2P system to facilitate resource discovery is presented.