Chapter 1

INTRODUCTION

The goal of this chapter is to highlight the challenges in mobile ad hoc networks raised by mobile computing and put the focus on new challenging problem, the service provisioning. To begin with, the mobile computing domain, its characteristics and constraints were examined. Service oriented architectures that use services as main building blocks were discussed. Then a discussion on the key research problems raised by enabling service provisioning in mobile ad hoc networks follows. Finally, the contributions followed by organization of the thesis were presented.

1.1 Motivation

1.1.1 Issues of Mobile Computing

Hardly a day passes without some new evidence of the proliferation of portable computers in the marketplace, or of the growing demand for wireless communication. Support for mobility has been the focus of number of experimental systems, research & commercial products since several decades. Although many basic principles of distributed system design described by George Coulouris et al (2001) continued to apply; mobile computing brought more constraints to research problems. Indeed the solution of many previously-encountered problems becomes more complex and new problems, related to mobility, arise.

In his article, Satyanarayanan (1996) listed four main constraints raised by mobile computing:

Resource-poor: mobile elements are resource-poor relative to static ones. Considerations of weight, power, size and ergonomics will cause a penalty in computational resources such as
processor speed, memory size and disk capacity. These characteristics are closely related to the embedded technologies of these elements. Even though mobile elements will improve as the technology is improving, they will always be resource-poor relative to static elements.

*Hazardous mobility:* mobile elements are more vulnerable to loss or damage. Their mobility makes it difficult to consider their availability. Security problem is also more important for mobile elements as their mobility lead them to many different places and situations, much more complex than in a static position.

*Variable connectivity:* mobile connectivity depends strongly on the mobile element’s geographical position. Some buildings may offer reliable, high-bandwidth wireless connectivity while others may only offer low-bandwidth connectivity. Outdoors, a mobile element may surely have to rely on a low-bandwidth wireless network with gaps in coverage.

*Finite energy resource:* while battery technology will undoubtedly improve over time, the problem of power consumption will not diminish. The finite energy resource is a fact that every developer of applications in mobile environments must deal with.

Due to these constraints, mobile computing raises several research problems listed by Mahadev Satyanarayana (2001). Many of the following problems of computer science deal with the challenges of mobile computing:

*Mobile networking:* protocols and techniques that allow mobile device users to move from one network to another by supporting routing to and from mobile hosts. Mobile networking includes Mobile IP defined by Pravin Bhagwat et al (1996), mobile ad hoc network protocols described by Elizabeth Royer and Chai-Keong Toh (1999) and techniques to improve TCP performance in wireless networks. This area of research tackles the network layer.

*Mobile information access:* in mobile environments, data and more generally information should be able to “follow” the mobile host. Mobile information access includes disconnected operation as per James J. Kistler and Mahadev Satyanarayanan (1992),

Adaptive protocols and applications: two levels of adaptation are required. Protocols have to adapt to a different set of parameters in mobile wireless networks. They need to be designed for adaptation to multiple parameters such as latency, burst error, disconnection during hand-off, asymmetry of the link, location and cost. The adaptation this thesis interested in is more about applications’ adaptation the concept described by Mahadev Satyanarayanan (1994). The support for adaptative applications includes adaptative resource management and context awareness.

QoS-aware system: with mobility, the path of information flow may change with each move; hence delays will be surely impacted by mobility. This area of research tries to adapt applications, services and all the functionalities provided to the clients to the ever changing world of mobile computing.

Location or position sensitivity: use of location, location sensing - the technique discussed by Roy Want et al (1992) and location-aware system behaviour are some of many research challenges for the mobile computing domain. These challenges related to location and positions are very often located in the network layer.

Though mobile computing brought challenges and constraints to distributed systems, it kept evolving with the evolution of the technology. The mission of mobile computing is to allow users to access any information using any device over any network at any time. As mobile ad hoc networks belong to mobile computing, these issues and research challenges are applicable to mobile ad hoc networks too.

1.1.2 Issues of Mobile Ad hoc Networks

Mobile ad hoc networks are collection of mobile nodes with no pre-established configuration or infrastructure. Any two devices can communicate with each other by the use of
short range wireless communication capabilities like Wi-Fi or Bluetooth interfaces, when they are in wireless range of one another. As per Perkins and Bhagwat (1994) mobile ad hoc networks are multi-hop networks where the wireless nodes communicate beyond their communication range by the use of intermediate nodes.

Since every node of a mobile ad hoc network is mobile and possibly very volatile they are having highly dynamic topology. These constant topological variations will eventually lead to a continuous state of network instability, which in turn can deteriorate the performance of services and applications on these networks. Another important issue is that typically devices participating in mobile ad hoc networks, have limited resources as far as storage and processing capabilities are concerned. By adding the energy constraints in the set of limitations on mobile ad hoc networks, one can easily understand that the road to fully grasping the potentials of mobile ad hoc networking is not going to be an easy one.

Mobile ad hoc network applications have been designed in the area of military and disaster relief. As the advancement in the wireless devices and the craze of using handhelds leads to the penetration of mobile ad hoc networks in the domains where all the handhelds, wireless communication and infrastructure free networks can be applied. However the highly dynamic nature of the topology, unpredictable conditions on the network connections, heterogeneous nature of the devices involved, constraints of the resources are the main hurdles to be removed for the wide spread application. Many works such as Marco conti and silvia Giordana (2007) and Sunil Kumar et al (2006) are mainly focusing on the lower layers of the mobile ad hoc networks where as the concentration on application layer issues are comparatively very less.

The applications based on mobile ad hoc networks must be self-adaptive due to continuous change in the execution environment. The changes in the environment like, node location, resource availability, network condition etc have their impact on the application. The primary issue among them is node mobility. The mobile nodes can leave network area without any notice, by suddenly disrupting the availability of local resources. Even worst, generally speaking mobility can lead to potential network partitions, due to the disconnection of traffic forwarders where all nodes belong to one network segment cannot communicate with nodes in the other.
Node symmetry and mobility arguments prevent once again the usage of fixed servers and called for highly distributed strategies. So, the applications have to be self-adaptive as far as possible at least they should be able to adapt the way they execute. As mobile ad hoc networks cannot assume the existence of any devices like GPS etc., this thesis uses the technique of publishing through advertisements the current status of the devices and their moving speed & direction and other parameters. The monitoring system based on event-driven programming is used to predict and fire events for appropriate actions in the system.

It is obvious there is a need for techniques that mitigate these problems in mobile ad hoc networks in order to be able to harness the vast range and diversity of enhancements and advantages that mobile ad hoc networking has to offer. In this respect context awareness can greatly assist in building and maintaining reliable, long term mobile ad hoc networks. Context awareness can assist in building more predictable and thus more reliable mobile ad hoc networks by providing higher and lower level information regarding the mobile node’s context, which can aid in predicting its mobility patterns. When the mobility of each and every node is predicted the future topology of the network can be foreseen, allowing thus the deployment of services and facilitating scalable routing protocols. The term context is used to indicate all the information accessible by the mobile node and descriptive of its surroundings whether they refer to computational or physical properties. The driving force behind this concept is based on the observation that if one has an idea of what is happening or what is going to happen in the mobile ad hoc network in the near future one can act proactively and adjust the mobile ad hoc network in order to preserve a certain degree of network stability.

As users move their computing environment moves with them, the availability of services changes depending upon the nodes in the network which may extend as users meet new nodes, or shrinks user node separate. Moreover users should be able to take full advantage of the local capabilities and resources within a given environment, even with its ever changing topology.

Undoubtedly, research in the mobile ad hoc networks has considerably advanced. At the networking level, for example, wireless communication is already possible through technologies like Bluetooth, Wi-Fi and Zigbee. Protocols which take care of both mobility and ad hoc networking have also been developed, like Mobile IP, GPRS, UPnP and Zeroconf. At an application level, when considering the context awareness of applications, undoubtedly some
efforts have been made in specific application domains. Applications of mobile ad hoc network may span to several domains such as home automation, health care, learning and multimedia. The task of building pervasive computing applications can be too tedious if performed from scratch. In other words, the developer will need to deal with low level networking protocols to high level applications context awareness. This, of course deviates the attention of the developers to tasks that are not the purpose of the application. Instead, they should only concentrate on the application logic, that is, the tasks the application must perform.

1.1.3 Service Oriented Architecture

Computing has become easy and useful as they are available in all the objects in our day-to-day life and they are capable of interacting with each other in a distributed way (Castelli (2007)). However in mobile environments the issue to be sort out for the development of framework to ensure software reusability is different than that of fixed wired networks. It is challenging because of the nodes involved in the network are mobile in nature and for the most importantly they are based on the radio communication and also multi-hop. This leads to the situation where the availability of the services and the working environment of the components vary from time to time.

Middleware/framework is software layer which supports designers building distributed mobile applications. Middleware aims at facilitating communication and coordination of distributed-components, concealing difficulties raised by mobility from application developers as much as possible. When developing distributed applications, designers do not have to deal explicitly with problems related to distribution, such as heterogeneity, scalability, resource sharing and the like. Middleware developed upon network operating systems provides application designers with a higher level of abstraction, hiding the complexity introduced by distribution and facilitate proper management of resources.

A distributed system consists of a collection of components, distributed over various computers. These components need to interact with each other in order to exchange data, or to
access each other’s services. A middleware is a layer between distributed system components and network operating system components and it facilitates the interaction between components.

There is a broad spectrum of potentially useful applications for mobile ad hoc networks, but application development in this domain is not easy. Obviously solving the same issues in every application from the scratch is not feasible. In software engineering authored by Somerville (1982), the concept of using the previously developed software modules for the new applications is effectively utilized under the term reuse. Creating reusable components is an art by itself. The Component based application development defined by Batista and Rodriguez (2000) has been evolved towards the realization of this software reuse. There were various paradigms defined ranging from Remote Procedure Calls (RPC) middleware as discussed in Judith (2002), the Message Oriented Middleware (MOM) middleware as discussed by Barb Gomolski (1997), the Object Request Broker (ORB) middleware defined by Arno Puder et al (2005) and the Service-Oriented Architectures (SOA) middleware were evolved.

This reuse of components can be realized by Service Oriented Architecture. It is the architecture which facilitates the publication of services and needed system can access them. This architecture fosters decoupling between interactive applicative entities. The main building blocks in SOA are services. Services are self-describing, open components that support rapid, low-cost development and deployment of distributed applications. The Service provisioning is a way of reusing the services defined already.

SOA is a powerful programming model. It provides a loose coupling of software components in order to facilitate the utilization of components defined by other organizations can also be used in the application development. It defines three entities Provider, Broker/Directory and Consumer. The principle of the SOA is a three step process, first step service publication by the provider and stored in the repository, the second one is the looking for required service in the repository and finally access the matched services. In other words SOA is a form of distributed system architecture that is typically characterised by the following properties as detailed by Nikola Milankovic (2006):

*Logical view:* the service is an abstracted, logical view of actual programs, databases, business processes and so on, defined in terms of what it provides.
**Message orientation:** the service is formally defined in terms of the messages exchanged between providers and requesters and not the properties of these latter themselves. The internal structure including features such as implementation language, process structure and even database structure, are deliberately abstracted away in the SOA: using the SOA one does not and should not need to know how a provider implements a service.

**Description orientation:** a service is described by machine processable meta-data. The description supports the public nature of the SOA: only those details that are exposed to the public and important for the use of the service should be included in the description.

**Granularity:** services tend to use a small number of operations with relatively large and complex messages.

**Network orientation:** services tend to be oriented toward use over a network, though this is not an absolute requirement.

**Platform neutral:** messages are sent in a platform-neutral, standardized format delivered through the interfaces. XML is usually the most obvious and used format that meets this constraint.

A computing infrastructure where everything is a service offers many new standards and application possibilities. Among the main challenges, however is the issue of standardized way of application development in such a heterogeneous environment. The natural way of doing this is by reusing the existing services and adapting the binding depending upon the mobility of the services transparently, as far as possible. In an open environment such as mobile ad hoc networks the ability of adapting to the mobility are only possible if services are implemented and described in interoperable languages and ability of knowing the current execution context.

In this thesis the concept of profile is used as meta data to define the Services - the software and hardware resource that can be shared. Service provisioning enables the provider entities to advertise and facilitate others to access the services. In the context of mobility, the
issue of service provisioning differs in many ways from the traditional way of service provisioning. The changes in standards as discussed by Wakefield et al (2007) & Jacqueline Floch (2006) and evolutions in the technology give raise to many new directions in the area of service provisioning.

1.2 Limitations

Using appropriate service provisioning framework in the mobile ad hoc environment is crucial to help the spreading of mobile ad hoc networks in everyday life. This framework must be able to cope with the inherent properties of Mobile ad hoc networks. The architectures of traditional service provisioning solutions used in communication and data networks (e.g., JINI (Sun Microsystems (2013), UPnP (UPnP Forum 2008)) are not well suited for mobile ad hoc networks and thus it is difficult to adapt them for the mobile ad hoc environment. Usually they are based on the centralized service management model using the client/server architecture. The shortcomings of the client/server architecture are the low fault tolerance due to the reliance on a single, central server node and the limited scalability of the server. So, fully distributed model using the peer-to-peer architecture is the only possibility to meet these shortcomings of having centralised server.

In traditional service discovery protocols, the focus is on capturing the static characterisation of services. They do not take into account the user’s current needs and situation or control policies. Service selection in mobile ad hoc networks not only based on the functional properties of the services they also depend on the non functional properties like cost, quality etc. These properties may also vary depending upon the user too. Therefore, service user should be provided with a mechanism to express their wish and preferences among the properties. Then only the selected service will fulfil the need of the user correctly. Guaranteeing more accuracy in the selection of available services is especially beneficial to mobile users, who are often on the move with constrained resources.

To increase the stability of the selected service provider node, taking node mobility and the battery power remaining into account is indispensible. High mobility of the node can result in
frequent changes of service provider node and thus the binding changes. Thus selecting the service providers based on their metrics which takes into consideration of moving speed and battery power remaining as metrics will lead to a stable provider selection. In this thesis a mechanism to calculate the metrics of the providers before assigning them to the client node have been developed.

The mobility of users, terminals and service components require novel solutions to handle a set of bindings to needed services. The bindings to the services must be properly arranged to maintain service accessibilities. The choice of the proper binding depends on several factors, from runtime conditions and access-device properties to security requirements and user preferences. Mobile applications require greater binding flexibility than is provided by conventional approaches, in which developers hard coded strategies into service logic. The flexibility in binding relieves from the requirement of generating multiple statically specified versions of an application to accommodate different scenarios. In this thesis a method for dynamic binding has been proposed by the use of policies specified by the application developers.

Adaptation is the process of making adjustments to suit the environment and to adjust different conditions. Adaptation is very important in dynamic and ever changing environments. The adaptation problem is one of the biggest challenges for software engineering (Jacqueline Floch (2006)). This is particularly the case since mobile computing which has tuned adaptation from the slow process of software evolution into a highly dynamic run-time procedure that needs to occur as devices and applications are on the move. There are different layer where adaptation can be applied: the service execution and deployment adaptation, the resource adaptation and so on. This thesis is interested in adaptation as a technique of service reselection and binding choice of services.

Next major step in service provisioning is maintenance. This step is not needed in the fixed network where the service provider will not move. Whereas this step essential in the case of mobile ad hoc networks in which node mobility is a norm. When something new happens in the environment it should have a means to identify them and adapt itself according to that. When some new services available which make the selection perfectly fit and provide best of what it
can do then that service has to be selected and provided to the user. When some services disappear from the environment the action to do is to adapt the environment to this loss. Good service selection localizes communication, which in turn reduces inter-node interferen-ces and allows for multiple concurrent retransmissions in different parts of the network. Less optimal service selection spreads traffic over the network, increases interferen-ces and reduces overall network throughput.

Over time, changes in network topology degrade the optimality of service selection requiring clients to continuously re-evaluate their choice of a server; this process is referred to as reselection. And even actively probe the network for availability of new service providers, this process is referred to as rediscovery.

Many research works are interested in one of these development trends, but none have tried to combine all these functionalities into one framework. Current works in the literature offers mechanisms to access context information, defining non functional properties and requirements. But there is no unified work which provides mechanism to access the current context, using context for best service selection, binding and rebinding decision in the field of service provisioning.

Motivated by the above observations, this thesis delves to find the required components and to propose a framework which enables the application development for mobile ad hoc networks by taking advantage of service oriented architecture for provisioning of services. In order to achieve this, following objectives have been formulated.

Centralized or semi centralized service discovery which are applied to fixed networks will not be suitable for mobile ad hoc networks, because of their assumption on the existence of the central entity to act as broker or coordinator. But the service requester must be served with a directory of services. This is also a question to be addressed. Mobile ad hoc network protocols are based on the lower layers of the ISO model whereas the service discovery is the issue at the application layer. The dynamic topology of the mobile ad hoc network nodes because of their mobility will affect the availability of the services. Moreover the dynamic nature and the current
execution context vary over time, mechanism has to be there to identify and handle the situation transparently without bothering the application as far as possible.

Different ways are there to represent the execution context; mechanism has to be devised to access them. Also each context may be of different data type so evaluation function for each of the data type has to be provided.

As there may be more than one provider in the network for the same service, mechanism has to be devised in order to allow the user to specify, how important the particular criteria and their preference over the choice to select the service.

1.3 Contributions of the Research

This thesis tackles above mentioned issues and proposes framework for service provisioning. First of all, it proposes a survey of available research work related to service provisioning in mobile ad hoc networks. Based on this survey following lacks were identified in service provisioning in mobile ad hoc networks: the unified vision of service selection and provision in the case of mobility, the functional and non functional description and usage for service selection, accessing the current user environment for service provisioning and spontaneous service reselection in case of service unavailability.

Access to remote services and related issues affect many applications in mobile ad hoc network and deserve to be jointly faced. The proposed framework operates at the application level mainly because several dissemination and retrieval decisions are possible only at this abstract layer. Thus, application developers only need to set service parameters, while the framework is transparently in charge of the actual operations. Finally working at application level also simplifies portability over heterogeneous communication technologies and routing protocols, by hiding low-layer implementation details from application developers. This particularly benefits mobile ad hoc network scenarios, where no networking standard has been developed and the stakeholders choice of action and accessing (providing) services may vary based on their context.
The primary idea behind the proposed approach is to improve service selection by current context of the client and the service provider. This thesis claims that designing schema for all the general-purpose cases will not be feasible. Therefore, this thesis focuses on a specific deployment scenario of increasing relevance for learning environment. These are limited spatial regions such as university campuses where many mobile wireless peers autonomously cooperate and maintain a node density almost invariant during a lone time interval.

A mechanism has been designed to identify and alter the service provider if needed based on the condition. To define when to switch over the service provider the proposed framework provided a means to define the event-condition-action in terms of policies. The monitors defined will identify the events and activate the actions accordingly. Three policies namely reselection, rediscovery and binding policies have been defined. Modules for policy decision and enforcement according to the policy specification were designed. To reduce the overhead and complexity only the stateless service provision is taken into consideration. The policy mechanism is also used to specify and actuate different bindings based on the needs without user interaction.

The framework proposes a Peer-to-Peer service cache at each node for each of the service available in the network. One group for each service type in the network is maintained and they are sorted based on the providers’ metric. In order to have up-to-date service availability details a set of protocol is designed and implemented.

Service discovery protocol has been designed and implemented, which uses cross layer information for its decision on the service selection. As the protocol resides at the application; cross-layer interactions layer is needed to take control on the topology as well as the current execution context.

Each context like user location, moving speed, bandwidth available etc. may be of different type. Mechanism to define different data typed context were provided. APIs for accessing the context values and evaluating based on their data type were developed. Modules
for evaluation and matching the non functional properties based on the preference levels were
designed and implemented.

Modules for the binding of the services dynamically were designed and implemented. The way how to react to different each contextual change may be provided by the developer of the application. Notification facility is also provided to intimate the application of its inability to satisfy the requirements, so that the application can take different actions.

Based on simulations each module of the proposed framework is compared with the existing work. From analysis it can be concluded that rebinding module is an important new module to be added to the service provisioning in mobile ad hoc networks. Context awareness makes better service selection. Making service provider’s selection based on their metric provides better performance than without metric selection.

1.4 Thesis Organization

The rest of the thesis is organized as follows. Chapter 2 provides the background information related to the research. Mobile ad hoc network and its characteristics, Service oriented architecture and technologies available to implement SOA have been listed. Overviews of the context-awareness, policy management were also presented. Chapter 3 provides an analysis of the requirements of a generic service provisioning framework for mobile ad hoc networks followed by the state of the art for service provision systems in mobile ad hoc networks and finally presents the existing service provision systems.

Chapter 4 introduces the design challenges. It shows the different phases of service provisioning and describes the integrated framework Architecture which takes into consideration of all these phases into consideration. It presents the challenging points that must be addressed when designing a service provision framework that remains viable in mobile ad hoc networks. Chapter 5 provides the distribution mechanism used to advertisement and how the queries are
matches. A modified form of service oriented architecture to cope with mobile ad hoc networks is defined and described.

Chapter 6 provides the description of non functional properties management. The influence of cross-layer interaction and how to access the dynamic context are discussed. The non functional property descriptions of a service and the evaluation are discussed. Each proposed work in this thesis is compared with similar contemporary work in terms of performance.

Chapter 7 presents the invocation and binding solutions. Invocations must be tolerant to communication delays due to disconnected network environments. The chapter discussed the binding strategies followed in the proposed work. The rediscovery and reselection policies for adaptation based on the context are described. The evaluation of these policies has been discussed. Chapter 8 concludes with the summary of the contributions and enumerates some of the future research directions.