CHAPTER 1

INTRODUCTION

Advances in telecommunication and the popularity of portable computing devices have facilitated the wireless ad hoc networks (WADNET) [1,2] to communicate between any pair of nodes at anytime and anywhere. This network can be further defined as infrastructureless type of wireless networks which does not rely on any central entity like routers or gateways to relay data. Due its uniqueness, it is extensively used in military tactical communications which is still considered as one of the primary application in WADNET [3]. Apart from tactical applications, WADNET have also drawn attention in commercial applications such as educational use, disaster recovery, and collaborative applications [3–5]. In all these applications, accessing data or sharing data to complete a mission becomes a critical issue due to recurrent changes in the network. This chapter discusses on types of WADNET, basic concepts of data replication, types of replication models which are suitable for WADNET and finally, confer the motivation that leads to the current contribution and the organization of the thesis.

1.1 OVERVIEW

WADNET are termed as the decentralized type of wireless network consisting of hand held devices that communicate in multi-hop fashion with each other as there is no pre-existing infrastructure when compared to other types of wireless networks [6, 7]. These networks [8] are generally classified as:

- Mobile ad hoc networks (MANET)
- Wireless Sensor Networks (WSN)
1.1.1 Mobile Ad hoc Networks

MANET is defined as a self-configuring infrastructureless networks comprising of mobile devices that communicate via wireless transmission [9–11]. It is classified as follows:

- **VANET**: It is defined as vehicular ad hoc network where communication take place between vehicles and roadside equipment.
- **InVANET**: Intelligence VANET is a kind of artificial intelligence that helps vehicle to behave in an intelligent manner during accidents, drunken-driving etc.
- **iManet**: Internet MANET are ad hoc networks that link mobile nodes and fixed Internet-gateway nodes to propagate data.

Typical applications of MANET include but not confined to military applications, collaborative and distributed computing and emergency and disaster recovery operations [12,13]. In all these applications, sharing data or information is essential with quick and reliable communication. However, the following characteristics make data accessibility a demanding problem.

- Topology is dynamic in terms of link quality and failure of nodes occurs due to energy depletion.
- As nodes are mobile, the topology changes occur frequently.
- The network is resource restraint in terms of bandwidth, energy and memory.

1.1.2 Wireless Sensor Networks

WSN are disseminated networks of small, lightweight wireless nodes deployed in large numbers to monitor the environment by measuring physical parameters such as temperature, pressure or relative humidity [14–16]. Sensor nodes are used in a variety of applications which require continuous monitoring and detection of specific events [17–19]. Initially, sensor nodes were used in military applications which include
battlefield surveillance, guidance of intelligent missiles and detection of mines. Soon after, the sensors were also extensively used in environmental applications such as forest fire detection, habitat monitoring weather forecasting etc.

In recent times, sensors are finding their own way in pervasive computing like remote health monitoring and also in commercial applications at home and industry. Hence, sensor network applications are continual and can be restricted only by human thoughts. The main goal of setting up a sensor network is to gather sensed data and propagate to a gateway node called as a sink. Accessing data from sensor networks is a challenging task due to the inherent nature of the network. The following are the inherent features of sensor networks that make data accessibility difficult:

- Sensor networks are mostly envisioned to be static but the network fluctuates in terms of link quality.
- Sensors rely only on battery power if it is deployed in unattended environments. Hence, depletion of energy may cause serious issues in gathering data.
- Sensor nodes are storage constraint, hence, additional storage resources are required to increase the capacity.

### 1.1.3 Similarities and Dissimilarities of MANET and WSN

WSN is similar to MANET in that both involve multi-hop communications [20, 21], both are said to be resource constrained networks and do not rely on base stations. However, the nature of the applications and routing requirements for the two are significantly different in several aspects which are as follows:

- The classic sort of communication in a sensor network is from multiple data sources to a data recipient / sink rather than communication between any pair of nodes.
- Since, the data being collected by multiple sensors is based on common event, it is liable to some redundancy in the data being communicated by the various
sources in sensor networks. Conversely, in MANET data shared by nodes may be distinct.

- Nodes interact with humans in MANET, whereas, nodes in sensor networks interact with the environment.
- Nodes are sparsely deployed in MANET, whereas, in sensor networks nodes are densely deployed.
- In many scenarios, the sensors are stationary particularly in unattended environment, while nodes participate in MANET are mobile. So, the nature of the dynamics in the two networks are different.
- Finally, the single foremost resource constraint in sensor networks is energy. The circumstances in sensor networks are much worse than that in MANET, where the communicating devices handled by human users can be replaced or re-energized fairly often. The scale of sensor networks and the necessity of unattended operation for months at a time, signify that, energy resources have to be handled even more circumspect.

1.1.4 Necessity for integration of MANET and WSN

Though these two networks have several dissimilarities, the integration of these two networks is a key element in the communications infrastructure. Both military and disaster-recovery operations will run more efficiently when people can communicate with each other as they move around, while at the same time getting up-to-date information from stationary wireless sensors describing the changes in the environment or hostile operations, where a human cannot be present always. The gathered information can be further shared with collaborators to complete an assignment. As these applications are time crucial, the effective data gathering in sensor network and sharing of data in a MANET is possible only by replicating data closer to recipients. Hence, to improve accessibility of data in MANET and WSN and to endure the dynamic changes in topology, the promising approach is to replicate data across multiple nodes in the network. The next section gives the need for replication and describes the types of replication models that are suitable for WADNET.
1.1.5 Data Replication

The fundamental thought in replication is to keep several copies or replicas of the same resources at various nodes in a network. A node can contact any of the nodes, preferably the closest replica, to access the shared data. Therefore, replication indeed helps to balance the load across nodes, reduces the access latency and avoid network congestion [22–24]. However, care should be taken in selecting best replica node so that the selected node is capable of serving multiple clients and can withstand the dynamic changes in the network.

1.1.6 Replication Models

Basically, replication in a network regardless of wired or wireless fit into any of three models named as Master Slave model, Client Server model and Peer to Peer model [24–26]. The details of these models are given below.

1.1.6.1 Master Slave Model

In this model, one copy of the replica is designated as the master and others are designated as slaves. The basic principle of this model is that the slave should always be identical to the master. The functionality of the slaves in this model is very limited, thus makes the design very simple. The slaves essentially are considered to be read-only. Most of the master-slave services ignore all the updates or modifications executed on the slave, and “undo” the update during synchronization, making the slave identical to the master. The modifications or the updates can be constantly performed only at the master and the slaves inturn synchronize directly with the master. Though, this model seems to be very simple in maintaining consistency, it is not suitable for distributed systems or mobile networks as it relies on a central entity, i.e, the master which may lead to a single point of failure.

1.1.6.2 Client Server Model

In the client server model, one replica node is chosen as special server and other nodes are chosen as clients. This model differs from master slave model by allowing updates to take place in clients but these updates should propagate to the server first
and are propagated further to all clients. The advantage of client server model is, it simplifies replication and reduces the message cost but suffers from slow propagation of updates and single point of failure (server failure).

Also, this model makes the functionality of clients very complex in a mobile environment. For instance, if client nodes move close to each other, they still cannot communicate directly to transfer the updates. Instead, the client has to upload the updates to a remote server that is far away from the client and other clients will download the updates only from the server. Hence, this indirect communication between clients leads to complexity and thus, limits the model in certain applications where the infra-structureless wireless environment and mobility of nodes are supported.

### 1.1.6.3 Peer to Peer Server Model

A different approach is presented in peer to peer model where every replica node has the same set of functionalities. In this model, nodes are not categorized as special nodes. Each node has equal responsibilities and they are termed as peers. Any replica in peer to peer model can make updates and propagate the updates directly to another replica node. The advantage of this model is that it increases the speed of the update propagation and decreases the frequency of using inconsistent data. Therefore, this model is well suitable for WADNET which do not require any pre-existing infrastructure and central administer.

### 1.1.7 Replication vs. Caching

Caching is defined as the process of pre-fetching the data and storing it in the cache memory of the client for future use. Whereas, replication is defined as the process of creating multiple copies of shared data and distributing to multiple servers closer to clients. Although, both mechanisms are intended to reduce the access delay, there exists a number of differences between them [27] which are listed as follows:

1. Replication is a proactive process that copies the data to selected servers. On the other hand, caching is the reactive process where needed or frequently used data is stored in local cache memory only if a client executes the query to the server to
fetch the data. Hence, replication is an autonomous process independent of query processing, whereas, caching is said to be a reliant process depending on query process.

2. In replication, replica remains in servers until they are explicitly deleted due to propagation of updates. On the other hand, in caching, data remain in cache until they are replaced by new client-interested data, by the replacement policy due to lack of space or due to invalidation of the data.

3. In replication, consistency of data is maintained by propagation protocols, whereas, in caching, invalidation protocols are used to make the cache copy invalidate or to remove out-of-date copies.

4. In replication, majority of shared data items is replicated across the existing servers, but in a caching small fraction of requisite data is copied in the cache.

Although, many caching schemes are proposed for MANET, its mere objective is only to reduce the data access delay. Hence, when compared to caching, the replication technique is considered to be more active owing to its high degree of performance. The performance of this technique is high due to subsequent reasons:

- Increasing the data availability and thus, making the network fault-tolerant.
- Splitting up the work among servers and balancing the load across the network.
- Making the network energy efficient by reducing the network traffic to access the data.
- Reducing the data access delay by placing the data server closer to clients.

1.2 MOTIVATION

MANET have a substantial deal over two decades since its establishment, predominantly in military scenarios where momentary network connections are requisite and time critical data sharing is essential for communication. This network consists of mobile nodes, which are capable of executing computations and data exchanges
among the peers. It helps to provide situational awareness among battle field entities like a dismounted platoon of soldiers, inter-battle field tankers and military aircraft. Situational awareness can be communicated in the form of a text or image or voice between the entities in the battlefield. Hence, information sharing in MANET becomes essential in a military environment to complete a certain mission.

Correspondingly, the considerable increase in wireless enabled portable devices have also triggered to configure MANET in new scenarios and provide services in limited spatial regions such as shopping malls, airport, meeting rooms etc. In all these regions, mobile peers have to collaborate with each other to share a common interest of data that are required for mobile users. Hence, MANET finds its use extensively both in tactical and commercial applications [28]. However, due to lack of static infrastructure, the network often gets partitioned, hence, data from one partition cannot be accessed by nodes in another partition. To alleviate this situation, the shared data must be replicated to ensure availability of data at all times.

Remarkably, research on MANET mainly concentrates only on developing and enhancing the existing routing protocols to adapt to dynamic environments [29–31]. Also, more attention is given to impose security mechanisms in MANET. However, very less attention is given on implementing middleware services where only very few state of the art proposals have addressed the challenging issues of data replication in MANET. Moreover, these proposed replication mechanisms are analyzed just by simulation and very little experimentation is done to implement them in real-time scenarios. Due to this reason, the number of applications in MANET are restricted both in military and commercial scenarios.

At present, data replication has gained its extensive attention in WSN as well. This type of network is known to be a special class of autonomous ad hoc networks that comprises of small and cheap sensors powered by low-energy batteries equipped with radio transceivers, and responsible for responding to physical parameters, such as temperature, pressure, magnetism and motion, by producing radio signals. These sensor devices have limited computing power, storage, and battery life and low transmission rates [32]. These networks are deployed for a specific application scenario to detect
some events.

At the outset, many proposed works in WSN are focused on node centric networking where sensed data from different sources are flooded toward the sink without knowing the sink’s data of interest. Two types of storage mechanisms are followed in node centric networks such as local storage (LS) and external storage (ES). In ES [33–35], nodes send data to the base station or sink as raw data and thus, generated traffic is highly directed from many nodes toward one or a few sink nodes creating a probable blockage. Besides, many nodes close to the sink become hot spot nodes. The extreme traffic, created due to the continuous reporting of sensor data, leads to higher consumption of energy per node, thus, reducing the overall lifetime of the network. In addition, ES approaches may also have an unbalanced energy consumption rate among nodes due to the variation of distance between the base station and sensor node. Furthermore, since the sink node is exclusively responsible to fuse data and answer all queries, ES mechanism may result to delayed service.

In LS scheme [36–39], each node keeps its sensed data in local memory. The query node in turn sends a blind query to retrieve the data of its interest since it does not know the target node. Due to these reasons, this mechanism also consumes more energy and incurs high network traffic. Therefore, the storage mechanisms discussed above have caused serious performance degradation by increasing the energy consumption, network congestion and end to end delay, consequently, leading to considerable packet loss. Hence, to improve the performance, data centric networking has become a paradigm shift from node centric networking.

The storage mechanism supported in data centric networking [17, 40–43] is termed as data centric storage (DCS). In this mechanism, a storage node exists between the source and sink. Hence, instead of flooding the query or raw data between the source and sink, if the sink is interested in a particular type of data it can retrieve the data from storage node rather querying the end source nodes. Though, this mechanism reduces the query cost, still it suffers from both high network communication cost and hot spot problem because there exist only a single storage node to store all the popular events. Thus, a different approach such as multi-replication data-centric storage (MR-DCS)
have been tried to address these limitations. MR-DCS [44, 45] achieves load balancing in storing events in numerous rendezvous nodes called as replication nodes or storage nodes, thereby bringing longer lifespan to sensor nodes.

Motivated by the facts discussed above, we identified that replication technique is considered to be a foremost and practicable solution both in the MANET and WSN to increase data availability, energy efficiency, and reduce the network traffic, thus, making the network resourceful. However, implementing replication in these types of networks becomes adverse owing to design challenges which are discussed in the next section.

1.3 REPLICATION DESIGN ISSUES

Data replication in WADNET must address additional new issues arising from the constraints imposed by the ad hoc network environment. The following are the issues which make data replication a challenging problem [27, 46, 47]:

• **Network partitioning issue:** Network disconnection occurs more repeatedly due to recurrent topology changes in ad hoc networks than in fixed networks. This issue scrupulously reduces data availability when the server that holds the preferred data is not in the same partition as the client nodes. Therefore, replicating data in future separate-partitions before the incidence of network partitioning can improve data availability. To accomplish this, the protocol should determine the time at which network partitioning might occur and replicate data items ahead of time.

• **Energy consumption issue:** Mobile nodes or Sensor nodes activate on low-power batteries. In a particular scenario, a single server may serve several clients, which causes its battery power to be depleted very rapidly. To improve data availability, the replication protocol should replicate the data items to new server before the node exhausts its battery power or reduce the communication cost to access the data to increase the lifetime of nodes. The new preferred server node should be competent enough to sustain for a long time period.
• **Scalability issue:** Many applications, particularly in sensor networks and tactical networks, require large size networks with tens of thousands of nodes. Increase of network size makes a query sent by a client node to navigate a lengthy path to reach the server node, thus increasing the query cost and latency. Moreover, the existence of a large number of querying nodes leads to more channel-access-contention among clients, which reduces significantly the available bandwidth and increases channel access delay. Furthermore, mobility of nodes may also affect the performance of the replication protocol. Hence, the replication protocol should be deliberate so that its performance is not very much affected if the number of nodes increases or the link quality of node decreases.

• **Real-time Applications:** Wireless ad hoc applications like rescue and military operations are time-critical where data have to be propagated in stipulated time without missing the time limit. Therefore, the replication technique should place the data closer to clients and be able to deliver correct information on or before the expiry of deadlines.

1.4 **SALIENT FEATURES OF DATA REPLICATION IN WIRELESS AD HOC NETWORKS**

To counteract the above-mentioned challenges, effective measures are mandatory to develop an efficient replication algorithm. The effectiveness of replication schemes is extremely sustained by the number of replicas to be created in the system, where to place the replica and when to relocate replica on the nodes by considering the stability of wireless links and other characteristics that exhibit the dynamic deeds [27]. The prominent features of replication mechanism in WADNET are as follows:

• The mechanism have to consider node mobility or link quality of nodes and migrate the data accordingly.

• Replication nodes ought to change over in time to balance the energy consumption.
• Number of messages to access the data must be reduced to improve energy efficiency.

• Data placement should always be closer to clients satisfying time critical applications.

• Server nodes must serve more number of clients which indirectly aid to increase the size of the network.

1.5 CONTRIBUTIONS AND ORGANIZATION

The main objective of this thesis is to propose a replication scheme that adapts to dynamic changes in topology, to improve the data availability, reducing the network traffic and minimize the query access delay by replicating data closest to clients.

The key contribution of this thesis is to select stable nodes as replica nodes both in the MANET and WSN. The node is said to be stable in MANET if does not move frequently from one location to another, whereas in WSN, if the network assumed to be static, the node is said to be stable only if it has better link quality with its neighbors and maximum energy to store the data. The selection process takes place in such a way that the placement of replicas should be closer to clients and the replica relocation ought to be dynamic to adapt the topology changes.

To determine the importance of nodes in a network for propagating data, we have identified appropriate centrality measures which are suitable for MANET and WSN. The originality of this contribution is that the centrality of a node is not decided only on topological connections. Additional characteristics are imposed to determine the centrality of a node. In MANET, we consider mobility as a major concern. Therefore, we introduce the mobility factor as a main metric to compute the centrality of a node. However, in WSN, particularly in stationary mode, the importance of a node is decided on three parameters such as link quality with its neighbors, residual energy of a node and the distance between the sink and the node.

Hence, two new centrality measures known as subgraph centrality for MANET and stability betweenness centrality for WSN are introduced to implement the replication
scheme. These centrality measures are computed using the weighted matrix as inputs which serve as an alternative of the adjacency matrix. To the best of our knowledge, our proposed work is the first work to introduce centrality measures for selection of replication points in WADNET.

1.5.1 Structure of thesis

The replication schemes that exist in the literature for both MANET and WSN are presented in detail in Chapter 2. The working mechanism and limitations of these replication schemes are discussed and analyzed. Finally, the need to develop a better replication scheme which can circumvent the challenges of WADNET is also discussed.

The proposed centrality based replication model for MANET is described in Chapter 3. The proposed mechanism is a two phased approach and each phase is explained by a set of algorithms. The performance of the proposed model is compared with the existing replication models and it shows that the proposed approach performs better than the existing state of the art mechanisms and it is well suited for dense and sparse networks.

Chapter 4 discuss the real time implementation of centrality based replication model with the support of OLSRd an extension of OLSR protocol in a wireless ad hoc testbed. The experimental results and evaluations show that the model performs well with a better response time and with increased probability of accessing the data in the minimum number of hops.

As data replication is considered to be more significant in WSN particularly for applications that are time critical, a replication scheme is proposed and its practicability is tested in a remote wireless sensor testbed. The implementation of the model is explained in detail in Chapter 5. The proposed model when compared to existing approaches shows better performance and proves that centrality measures aid to identify stable nodes to replicate data.

In Chapter 6, we conclude by summarizing the proposed work and pointing to future research directions in this work.