CHAPTER 2

A SURVEY OF ENERGY CONSERVATION SCHEMES IN WIRELESS SENSOR NETWORKS

2.1 INTRODUCTION TO WSN

A wireless sensor network consists of sensor nodes deployed over a geographical area for monitoring physical phenomena like temperature, humidity, vibrations, seismic events, and so on. Typically, a sensor node is a minute device that consists of three components such as a sensing subsystem for data attainment from the physical surrounding environment, a processing subsystem for local data processing and storage, and a wireless communication subsystem for data transmission. In addition, an energy source supplies the energy needed by the device to perform the planned task. Energy consumption is one of the biggest constraints of the wireless sensor node and this limitation combined with a typical deployment of large number of nodes has added many challenges to the design and management of wireless sensor networks.

They are typically used for remote environment monitoring in areas where providing electrical power is difficult. Therefore, the devices need to be powered by batteries and alternative energy sources discussed by Sandra Sendra et al (2011). Because battery energy is limited, the use of different techniques for energy saving is one of the hottest topics in WSNs. This energy source habitually consists of a battery with a limited energy resource. In
addition, it could be inconvenient to recharge the battery, because nodes may be deployed in a hostile or unpractical environment.

On the other hand, the sensor network should have a lifetime long enough to fulfill the application requirements. In many cases, a lifetime in the order of several months, or even years, may be required. In some cases, it is possible to scavenge energy from the external environment (e.g. by using solar cells as energy source). However, external energy supply sources often display a non-continuous behavior so that an energy buffer (a battery) is needed as well. In any case, energy is a very critical resource and must be used very sparingly. Therefore, energy conservation is a key issue in the design of systems based on wireless sensor networks.

The main components of a sensor node as seen from the Figure 2.1 are microcontroller, transceiver, external memory, energy source and one or more sensors. Microcontroller performs tasks, processes data and controls the functionality of other components in the sensor node. Sensors are used to sense the data from the physical environment, memory is for storage, and a transceiver is used for data transmission.
Experimental results have shown that generally data transmission is very costly in terms of energy consumption, whereas consumption by data processing is considerably less. (Raghunathan et al 2002). The energy cost of transmitting a single bit of information is approximately the same as that needed for processing a thousand operations in a typical sensor node. (Pottie et al 2000). Energy consumption of the sensing subsystem depends on the specific sensor type. In many cases it is negligible with respect to the energy consumed by the processing and, above all, the communication subsystems.

An application can be characterized as one of the following three applications: a regular application, an application requiring a high communication rate, and a burst rate based application discussed by Ramassamy et al (2012). In other cases, the energy expenditure for data sensing may be comparable to or even greater than the energy needed for data transmission. In general, energy-saving techniques focus on two subsystems: the networking subsystem (i.e.energy management is taken into account in the operations of each single node, as well as in the design of networking protocols) and the sensing subsystem (i.e.techniques are used to reduce the amount or frequency of energy-expensive samples). In this chapter, the techniques used for energy conservation in wireless sensor networks have been surveyed Anastasi et al (2009). In many existing protocols for energy – efficient routing them forward packets through minimum energy path to the sink towards the minimum energy consumption, which cause an unbalanced distribution of residual toward all sensor node in network partitioned discussed by Jiwan et al (2013).
2.2. ENERGY CONSERVATION SCHEMES

Sensor networks are deployed in an ad hoc fashion, with individual nodes remaining largely inactive for long periods of time, but then becoming suddenly active when something is detected. Sensor Networks are generally battery constrained. They are prone to failure, and therefore the sensor network topology changes frequently discussed by Sinha, J.D et al (2012). Prior to discussing the advanced classification of energy conservation proposals, it is very significant to present the network-level and node-level architectures. Obviously, the energy breakdown heavily depends on the specific node as given below:

- The communication subsystem has much higher energy consumption than the computation subsystem. It has been shown that transmitting one bit may consume as much as executing a few thousands instructions. Therefore, communication should be traded for computation.

- The radio energy utilization is of the same order in the reception, transmission, and idle states, while the energy spending drops off at least one order of magnitude in the sleep state. Therefore, the radio should be put to sleep (or turned off) whenever possible.

- Depending on the specific application, the sensing subsystem might be another major source of energy consumption, so its energy consumption has to be reduced as well.

Based on the above architecture and energy breakdown, numerous approaches have to be exploited, at the same time, to reduce energy
consumption in wireless sensor networks. At a very common level, three main enabling techniques, namely duty cycling, data-driven approaches and mobility are identified.

**Energy conservation Schemes**

- **Duty cycling**
  - Topology Control
    - Location-driven
    - Connection-driven
  - Sleep/wakeup protocols
    - On-demand
    - Scheduled rendezvous
    - Asynchronous
  - MAC (Media Access Control) protocols with low duty cycle
    - TDMA (Time Division Multiple Access)
    - Contention based
    - Hybrid

- **Data-driven**
  - Data reduction
    - In-network processing
    - Data compression
    - Data prediction
Duty cycling is primarily focused on the networking subsystem. The most efficient energy-conserving operation is putting the radio transceiver in the (low-energy) sleep mode whenever communication is not required. Ideally, the radio should be switched off as soon as there is no more data to send/receive, and it should be resumed as soon as a new data packet is prepared. In this way, nodes interchange between active and sleep periods depending on network activity. This activity is generally referred to as duty cycling and duty cycle is defined as the part of time when nodes are active during their lifetime.

A common approach for conserving channel capacity and energy is optimizing power-aware routing and different kind of duty cycling (DC) and harvesting technology discussed by Glatz et al (2011). As the sensor nodes
execute a cooperative task, they need to coordinate their sleep/wake up times. A sleep/wake up scheduling algorithm thus accompanies any duty cycling scheme. It is typically a distributed algorithm based on which sensor nodes decide when to change from active to sleep and back. It allows neighbouring nodes to be active at the same time, thus making packet exchange possible even when nodes operate with a low duty cycle (i.e. they sleep for most of the time).

Duty-cycling schemes are normally unaware of data that are sampled by sensor nodes. In fact, data sensing impacts on sensor nodes' energy consumption in two ways:

- **Extra samples.** As the sampled data generally has strong spatial and/or temporal relationship (Vuran et al, 2006), there is no need to communicate the redundant information to the sink.

- **Energy consumption of the sensing subsystem.** Reducing communication is not enough when the sensor itself is energy starving.

In the first case, extra samples result in useless energy consumption. Even if the cost of sampling is small, it results in extra communications. The second problem arises every time the consumption of the sensing subsystem is not small.

2.3.1 **Topology Control Protocols**

The idea of topology control is strictly related to that of network redundancy. Dense sensor networks usually have some degree of redundancy. Energy conservation is a very critical issue in WSN. A lot of work has been
done on the techniques of topology control so that sensor nodes which are not in direct use can be put to a low power consuming state, thus saving energy discussed by Banerjee et al (2013). In many cases, network deployment is done at random, e.g. by dropping a large number of sensor nodes from an airplane. Therefore, it may be convenient to deploy a number of nodes greater than necessary to cope up with the possible node failures occurring during or after the deployment. In various contexts, it is much easier to deploy at first a greater number of nodes than re-deploying additional nodes when needed. Ganesan et al (2004) presented in a paper that the similar basis, a redundant deployment may be convenient even when nodes are placed by hand.

A number of criteria can be used to choose which nodes are to be activated/deactivated and when. In this view, topology control protocols can be broadly classified into two categories. One is location driven protocols that describe which node should be turned on and when, based on the location of sensor nodes which are assumed to be known. Another is the connectivity driven protocols which dynamically activate/deactivate sensor nodes in such a way that network connectivity or complete sensing coverage (Kong et al 2007) is fulfilled.

2.3.1.1 Location-driven

GAF (Geographical Adaptive Fidelity) is a location-driven protocol that reduces energy consumption while keeping a constant level of routing fidelity. The sensing area where nodes are distributed is divided into small virtual grids. All nodes within the same virtual grid are equivalent for routing and just one node at a time needs to be active. Therefore, nodes have to coordinate with each other to decide which one can sleep and how long.

Initially, a node starts in the discovery state where it exchanges discovery messages with other nodes. After broadcasting the message, the
node enters the active state. While active, it periodically re-broadcasts its discovery message. A node in the discovery or active state can change its state to sleep when it detects that some other equivalent node will handle routing. Nodes in the sleeping state wake up after a sleeping time and go back to the discovery state. In GAF, load balancing is achieved through a periodic re-election of the leader, i.e. the node will remain active to manage routing in the virtual grid. The leader is chosen through a rank-based election algorithm which considers the nodes' residual energy, thus allowing the network lifetime to increase in proportion to node density.

GAF is independent of the routing protocol, so that it can be used along with any existing solution of that kind. In addition, GAF does not significantly affect the performance of the routing protocol in terms of packet loss and message latency. However, the structure imposed over the network may lead to an underutilization of the radio coverage areas. In fact, as all nodes within a virtual grid must be able to reach any node in an adjacent virtual grid, the nodes are actually forced to cover less than half the distance allowed by the radio range.

Although being defined as a geographic routing protocol, GeRaF (Geographic Random Forwarding) actually presents features which are in the direction of location-driven duty-cycled operations. They make use of both nodes position and redundancy. Nodes follow a given duty cycle to switch between awake (active) and sleep (inactive) states. Nodes periodically switch to the active state, starting with a listening time, so that they can participate in routing if needed. Data forwarding starts as soon as a node has a packet to send. In this case, the node becomes active and broadcasts a packet containing its own location and the location of the intended receiver. Then a receiver-initiated forwarding phase takes place. As a result, one of the active
neighbors of the sender will be selected to relay the packet towards the destination.

The main idea is that each active node has a priority which depends on its closeness to the intended destination of the packet. In addition to priority, a distributed randomization scheme is also used, in order to reduce the probability that many neighboring nodes are simultaneously sleeping. Specifically, the portion of the coverage area of the sender which is closer to the intended destination is split into a number of regions. Each region has its associated priority and regions are chosen so that all the nodes within a region are closer to the destination than any other node in a region with a lower priority.

After the broadcast, nodes in the region with the higher priority contend for forwarding. If only one node gets the channel, it simply forwards the packet and the process ends. Otherwise, multiple nodes may transmit simultaneously, resulting in a collision. In this case, a resolution technique is applied in order to select a single forwarder. There may also be the case in which no node can forward the packet because all nodes in the region are sleeping. In the next transmission attempt, the forwarder will be chosen among nodes in the second highest-priority region and so on. Every time the relay selection phase will be repeated until a maximum number of retries will be reached. Eventually, after a hop-by-hop forwarding, the packet will reach the intended destination.

2.3.1.2 Connectivity-driven

Span is a connectivity-driven protocol that adaptively elects "coordinators" of all nodes in the network. Coordinators stay awake continuously and perform multi-hop routing, while the other nodes stay in sleeping mode and periodically check if it is needed to wake up and become
a coordinator. To guarantee a sufficient number of coordinators Span
uses the following coordinator eligibility rule: if two neighbors of a
non-coordinator node cannot reach each other, either directly or via one or
more coordinators, that node should become a coordinator. However, it may
happen that several nodes discover the lack of a coordinator at the same time
and thus, they all decide to become a coordinator. To avoid such cases, nodes
that decide to become a coordinator defer their announcement by a random
backoff delay.

Each node uses a function that generates a random time by taking
into account both the number of neighbors that can be connected by a
potential coordinator node and its residual energy. The fundamental ideas are
that (i) nodes with a higher expected lifetime should be more likely to
volunteer to become a coordinator and (ii) coordinators should be selected in
such a way as to minimize their number. Each coordinator periodically
checks if it can stop being a coordinator. In detail, a node should withdraw as
a coordinator if every pair of its neighbors can communicate directly or
through some other coordinators.

To avoid loss of connectivity, during the transient phase, the old
coordinator continues its service until the new one is available. The Span
election algorithm requires knowing the neighbor and connectivity
information to decide whether a node should become a coordinator or not.
Such information is provided by the routing protocol. Hence span depends on
it and it may require modification in the routing look up process.

ASCENT (Adaptive Self-Configuring SEnsor Networks Topologies) is
a connectivity-driven protocol that, unlike span, does not depend on the
routing protocol. In ASCENT, a node decides whether to join the network or
continue to sleep based on information about connectivity and packet loss
that are measured locally by the node itself. The basic idea of ASCENT is
that initially only some nodes are active, while all other ones are passive, i.e. they listen to packets but do not transmit. If the number of active nodes is not large enough, the sink node may experience a high message loss from sources.

The sink then starts sending help messages to solicit neighboring nodes that are in the passive state (passive neighbors) to join the network by changing their state from passive to active (active neighbors). Passive neighbors have their radio on and listen to all packets transmitted by their active neighbors. However, they do not cooperate to forward data packets or exchange routing control information they only collect information about the network status without interfering with other nodes. On the contrary, active neighbors forward data and routing (control) messages until they run out of energy.

Active nodes can also send help messages when they find the local data loss at an unacceptable level. As soon as it joins the network, a node starts monitoring the network conditions and also signals its presence as an active node through a neighbor announcement message. This process continues until the number of active nodes is such that the message loss rate experienced by the sink is below a pre-defined application-dependent threshold.

The process will re-start when some future network event (e.g. a node failure) or a change in the environmental conditions causes an increase in the message loss. As mentioned above, ASCENT is independent of the routing protocol. In addition, it limits the packet loss due to collisions because the nodes density is explicitly taken into account as a parameter (in the form of a neighbor threshold value).
2.3.2 Sleep/wake up Protocols

As previously discussed, sleep/wake up schemes can be defined for a given component (i.e. the radio subsystem) of the sensor node, without relying on topology or connectivity aspects. In this section, a survey of the main sleep/wake up schemes which are implemented as independent protocols on top of the MAC protocol (i.e. at the network or the application layer) has been carried out. Independent sleep/wake up protocols can be further subdivided into three main categories: on-demand scheduled rendezvous and asynchronous schemes.

2.3.2.1 On-demand

On-demand protocols take the most intuitive approach to energy management. The basic idea is that a node should wake up only when another node wants to communicate with it. The main problem associated with on-demand schemes is how to inform the sleeping node that some other node is willing to communicate with it.

Actually, the approach taken by on-demand protocols is the ideal one, because it maximizes energy saving as nodes remain active only for the minimum time required for communication. In addition, there is only a very limited impact on latency because the target node wakes up immediately as soon as it realizes that there is a pending message. Unfortunately, the adoption of a radio triggered wake up scheme is almost always impractical, since it can only be applied when the distance between nodes is very short indeed (a few meters). Hence, introducing an additional wakeup radio is a more promising direction, especially suitable to event detection applications.

However, the wake up radio is costly and generally it is not shipped with commonly used sensor platforms. So when a second radio is not
available or convenient, other solutions such as the scheduled rendezvous and the asynchronous wake up schemes can be used. Both of them trade energy savings for an increased latency experienced by messages to travel through several hops.

2.3.2.2 Scheduled rendezvous

An alternative solution is using a scheduled rendezvous approach. The basic idea behind scheduled rendezvous schemes is that each node should wake up at the same time as its neighbours. Typically, nodes wake up according to a wake up schedule and remain active for a short time interval to communicate with their neighbours. Then, they go to sleep until the next rendezvous time.

The scheduled rendezvous approach is convenient because it is suitable to data aggregation and supports broadcast traffic. Unfortunately, it requires nodes to be synchronized, in which some cases can be difficult to achieve or expensive in terms of additional protocol overhead for synchronization. On the other hand, asynchronous wake up protocols do not need a tight synchronization among network nodes. In addition, asynchronous schemes are generally easier to implement and can ensure network connectivity even in highly dynamic scenarios where synchronous (i.e. scheduled rendezvous) schemes become inadequate. This greater flexibility is compensated by lower energy efficiency.

2.3.2.3 Asynchronous

Finally, an asynchronous sleep/wake up protocol can be used. With these protocols, a node can wake up when it wants and still be able to communicate with its neighbours. This goal is achieved by properties implied
in the sleep/wake up scheme. Thus no explicit information exchange is needed among the nodes.

In the asynchronous schemes, nodes need to wake up more frequently than in scheduled rendezvous protocols. Therefore, asynchronous protocols usually result in a higher duty cycle for network nodes than their synchronous counterparts. In addition, the support to broadcast traffic is problematic. Due to their wider applicability and their properties, scheduled rendezvous and asynchronous approaches seem to be the most promising solutions in the class of sleep/wake up protocols.

2.3.3 MAC Protocols with Low Duty Cycle

Several MAC protocols for wireless sensor networks have been proposed and many surveys and introductory papers on MAC protocols are available in the literature review discussed by Demirkol et al (2006), Langendoen et al (2008), Naik et al (2004), and Heidemann et al (2004). In the following discussion, focus is mainly on energy management issues rather than on channel access methods. However, most of them implement a low duty-cycle scheme for energy management.

The tremendous and rapid development in sensors technology allowed their application in various fields requiring monitoring, such as, transportations, rare species surveillance, agriculture, military activities, medical field, etc. Due to their intrinsic constraints and limitations, several dedicated MAC protocols have been designed for wireless sensor networks and whose main objectives are bandwidth optimization while keeping very low energy consumption discussed by Sabri Khssibi et al (2013).

In the physical layer, the use of directive and adaptive phased arrays is proposed for the WSN gateways, increasing the communication range
between sensors and their gateway discussed by Chih-Kuang Lin et al (2013). In wireless communication, idle listening, receiving and transmitting are the main source of consumption of energy discussed by Alshaibi et al (2012). The most common MAC protocols are TDMA (Time Division Multiple Access) based, contention-based, and hybrid.

2.3.3.1 TDMA

TDMA (Time Division Multiple Access) schemes naturally enable a duty cycle on sensor nodes as channel access is done on a slot-by-slot basis. As nodes need to turn on their radio only during their own slots, the energy consumption is ideally reduced to the minimum level required for transmitting/receiving data. In TDMA-based MAC protocols, time is divided into (periodic) frames and each frame consists of a certain number of time slots. Every node is assigned to one or more slots per frame, according to a certain scheduling algorithm, and it uses such slots for transmitting/receiving packets to/from other nodes.

In many cases, nodes are grouped to form clusters with a cluster-head which is in charge of assigning slots to nodes in the cluster, LEACH and Energy-aware TDMA-based MAC. One of the most important energy efficient TDMA protocol for wireless sensor networks is TRAMA. TRAMA divides time into two portions, a random-access period and a scheduled access period. The random access period is devoted to slot reservation and is accessed with a contention-based protocol. On the contrary, the scheduled access period is formed by a number of slots assigned to an individual node.

The slot reservation algorithm is as follows. First, nodes that derive two-hop neighborhood information are required to establish collision free schedules. Then, nodes start an election procedure to associate with each slot having a single node. Every node gets the priority of being the owner of a
specific slot. This priority is calculated as a hash function of the node identifier and the slot number. The node with the highest priority becomes the owner of a given slot.

Finally, nodes send out a synch packet containing a list of intended neighbor destinations for subsequent transmissions. As a consequence, nodes can agree on the slots which they must be awake in. Unused slots can be advertised by their owners for being re-used by other nodes.

FLAMA (FLow-Aware Medium Access) is a TDMA MAC protocol derived from TRAMA and optimized for periodic monitoring applications. The main idea is to avoid the overhead associated with the exchange of traffic information. As the message flow in periodic reporting applications is rather stable, FLAMA first sets up flows and then uses a pull-based mechanism, so that the data are transferred only after being explicitly requested.

2.3.3.2 Contention-based protocols

Contention-based protocols are the most popular class of MAC protocols for wireless sensor networks. They achieve duty cycling by tightly integrating channel access functionalities with a sleep/wake up scheme similar to those described above. The only difference is that in this case the sleep/wake up algorithm is not a protocol independent of the MAC protocol, but is tightly coupled with it.

Most of MAC protocols proposed for wireless sensor networks are contention-based protocols. One of the most popular contention-based MAC protocols is B-MAC (Berkeley MAC), a low complexity and low power MAC protocol which is shipped with the TinyOS operating system. The goal of B-MAC is to provide a few core functionalities and an energy efficient
mechanism for channel access. First, B-MAC implements basic channel access control features: a backoff scheme, an accurate channel estimation facility and optional acknowledgements.

Second, to achieve a low duty cycle B-MAC uses an asynchronous sleep/wake scheme based on periodic listening called Low Power Listening (LPL). Nodes periodically wake up to check the channel for activity. The period between consecutive wakeups is called check interval. After waking up, nodes remain active for a wake up time, in order to properly detect eventual ongoing transmissions. While the wake up time is fixed, the check interval can be specified by the application. B-MAC packets are made up of a long preamble and a payload. The preamble duration is at least equal to the check interval so that each node can always detect an ongoing transmission during its check interval. This approach does not require nodes to be synchronized. In fact, when a node detects channel activity, it just remains active and receives first the preamble and then the payload.

A well-known MAC protocol for multi-hop sensor networks is S-MAC (Sensor-MAC), which adopts a scheduled rendezvous communication scheme. Nodes exchange sync packets to coordinate their sleep/wake up periods. Every node can establish its own schedule or follow the schedule of a neighbor by means of a random distributed algorithm. Nodes using the same schedule form a virtual cluster. A node can eventually follow both schedules if they do not overlap, so that it can bridge communication between different virtual clusters.

The channel access time is split into two parts. In the ‘listen’ period nodes exchange sync packets and special control packets for collision avoidance. In the remaining period, the actual data transfer takes place. The sender and the destination node are awake and talk to each other. Nodes not concerned with the communication process can sleep until the
next listen period. To avoid high latencies in multi-hop environments S-MAC uses an adaptive listening scheme. A node overhearing its neighbor's transmissions wakes up at the end of the transmission for a short period of time. If the node is the next hop of the transmitter, the neighbor can send the packet to it without waiting for the next schedule. The parameters of the protocol, i.e. the ‘listen and the sleep’ period, are constants and cannot be varied after the deployment.

2.3.3.3 Hybrid protocols

Hybrid protocols adapt the protocol behavior to the level of contention in the network. They behave as contention-based protocols when the level of contention is low and switch to a TDMA scheme when the level of contention is high. The basic idea behind hybrid MAC protocols is switching the protocol behavior between TDMA and CSMA, depends on the level of contention and this is not new.

In PTDMA, time is slotted and nodes are distinguished as owners and non-owners. The protocol adjusts the access probability of owners and non-owners depending on the number of senders. By doing so, it adapts the MAC protocol to work as a TDMA or CSMA scheme depending on the level of contention in the network. However, PTDMA was conceived for a one-hop wireless scenario. Therefore, it does not take into account issues such as topology changes, synchronization errors, interference irregularities which are quite common in wireless sensor networks.

In the specific context of wireless sensor networks, one of the most interesting hybrid protocols is Z-MAC. In order to define the main transmission control scheme, Z-MAC starts a preliminary set up phase. By means of the neighbor’s discovery process each node builds a list of two-hop neighbors. Then a distributed slot assignment algorithm is applied to ensure
that any two nodes in the two-hop neighborhood are not assigned to the same slot. As a result, it is guaranteed that no transmission from a node to any of its one-hop neighbor interferes with any transmission from its two-hop neighbors.

The local frame exchange is aimed at deciding the time frame. Z-MAC does not use a global frame equal to all nodes in the network. It would be very difficult and expensive to adapt when a topology change occurs. Instead, Z-MAC allows each node to maintain its own local time frame that depends on the number of neighbors and avoids any conflict with its contending neighbors. The local slot assignment and time frame of each node are then forwarded to its two-hop neighbors. Thus any node has slot and frame information about any two-hop neighbors and all these synchronize to a common reference slot.

2.4 DATA-DRIVEN APPROACHES

Data driven techniques presented below are designed to reduce the amount of sampled data by keeping the sensing precision within an acceptable level for the application. Data-driven approaches can be separated according to the difficulty they address. In particular, data-reduction schemes address the case of unwanted samples, as energy-efficient data acquisition schemes are mainly aimed at reducing the energy spent by the sensing subsystem.

2.4.1 Data Reduction

It is important to discuss here one more classification level related to data-reduction schemes. These techniques are in-network processing, data compression and data prediction. All these techniques aim at reducing the quantity of data to be delivered to the sink node. Scenarios where nodes have limited energy and forward messages of different priorities are frequent in the
context of wireless sensor networks. Tailored to those scenarios, this relies on stochastic tools to develop selective message forwarding schemes discussed by Arroyo-Valles et al (2011).

2.4.1.1 In-network processing

In-network processing consists of performing data aggregation (e.g. computing average of some values) at intermediate nodes between the sources and the sink. In this way, the amount of data is reduced while traversing the network towards the sink. The most suitable in-network processing technique depends on the specific application and it must be modified accordingly.

2.4.1.2 Data Compression

Data compression can be applied to reduce the amount of information sent by source nodes. This scheme involves encoding information at nodes which generate data and decode it at the sink. There are different methods available for data compression which Pradhan et al (2003), Tang et al (2004), Wu et al (2006) and Xiong et al (2004) discussed in the papers.

2.4.1.3 Data Prediction

Data prediction involves building an abstraction of a sensed phenomenon, i.e. a model describing data evolution. The model can expect the values sensed by sensor nodes within certain error limits and they reside both at the sensors and at the sink. If the required accuracy is fulfilled, queries issued by users can be evaluated at the sink through the model without the need to get the accurate data from the nodes.
There are two instances of a model in the network, one residing at the sink and the other at source nodes (so that there are as many pairs of models as sources). The model at the sink can be used to answer queries without requiring any communication, thus reducing the energy consumption. Clearly, this operation can be performed only if the model is a valid representation of the phenomenon at a given instant.

Sensor nodes take data samples as usual and compare the actual data against the prediction. If the sensed value falls within an application-dependent tolerance, then the model is considered valid. Otherwise, the source node may transmit the sampled data and/or start a model update procedure involving the sink as well. The features of a specific data prediction technique depend on the way the model is built. Data prediction techniques can be split into three main classes:

(i) Stochastic approaches
(ii) Time series forecasting
(iii) Algorithmic approaches

Techniques belonging to the first class derive a stochastic characterization of the phenomenon, i.e. in terms of probabilities and/or statistical properties. Two main approaches of this kind are as follows. On the one hand, it is possible to map data into a random process described in terms of a probability density function. Data prediction is then obtained by combining the computed probability density function with the observed samples. On the other hand, a state space representation of the phenomenon can be derived, such that forthcoming samples can be guessed by filtering out a non-predictable component modeled as noise.
The second class of data prediction techniques is time series forecasting, where a set of historical values (the time series) obtained by periodical samplings is used to predict a future value in the same series. Generally, a time series can be represented as a combination of a pattern and a random error. The pattern, in turn, is characterized by its trend, i.e. its long-term variation and its seasonality, i.e. its periodical fluctuation.

Finally, the last class of data prediction techniques relies on a heuristic or a state-transition model describing the sensed phenomenon. Such algorithmic approaches derive methods or procedures to build and update the model on the basis of the chosen characterization.

### 2.4.2 Energy Efficient Data Acquisition

An emerging class of applications is actually sensing-constrained. This is in contrast with the general assumption that sensing is not relevant from an energy consumption standpoint. In fact, the energy consumption of the sensing subsystem not only may be relevant, but it can also be greater than the energy consumption of the radio or even greater than the energy consumption of the rest of the sensor nodes (Alippi et. al. 2007). This can be due to many different factors as given below:

- Energy hungry transducers. Some sensors intrinsically require high energy resources to perform their sampling task.


- Active sensors. Another class of sensors can get data about
the sensed phenomenon by using active transducers (e.g. sonar, radar or laser rangers). In this case, sensors have to send out a probing signal in order to acquire information about the observed quantity, as discussed by Ditzel et al (2006).

- Long acquisition time. The acquisition time may be in the order of hundreds of milliseconds or even seconds; hence the energy consumed by the sensing subsystem may be high, even if the sensor energy consumption is moderate.

In this case, reducing communications may not be enough, but energy conservation schemes have to actually reduce the number of acquisitions (i.e. data samples). It should also be pointed out that energy efficient data acquisition techniques are not exclusively aimed at reducing the energy consumption of the sensing subsystem. By reducing the data sampled by source nodes, they decrease the number of communications as well. Actually, many energy-efficient data-acquisition techniques have been conceived for minimizing the radio energy consumption, under the assumption that the sensor consumption is negligible.

The classification of approaches for energy-efficient data acquisition presented by Raghunathan et al (2006) is as follows:

(i) Adaptive sampling

(ii) Hierarchical sampling

(iii) Model-based active sampling
2.4.2.1 Adaptive sampling

Adaptive sampling techniques exploit such similarities to reduce the amount of data to be acquired from the transducer.

2.4.2.2 Hierarchical sampling

The hierarchical sampling approach assumes that nodes are equipped with different types of sensors. As each sensor is characterized by a given resolution and its associated energy consumption, this technique dynamically selects which class needs be activated, in order to get a tradeoff between accuracy and energy conservation.

2.4.2.3 Model–based Active sampling

Model-based active sampling takes an approach similar to data prediction. A model of the sensed phenomenon is built upon sampled data, so that future values can be forecast with certain accuracy.

2.5 MOBILITY-BASED SCHEMES

In case, some of the sensor nodes are mobile, mobility can finally be used as a tool for reducing energy consumption (beyond duty cycling and data-driven techniques). In a static sensor network, packets coming from sensor nodes track a multi-hop path towards the sink (s). Thus, a few paths can be more loaded than others and nodes closer to the sink have to relay more packets so that they are more subject to early energy reduction as given by Mohapatra et al (2007). If some of the nodes (including, possibly, the sink) are mobile, the traffic flow can be changed if mobile devices are responsible for data collection directly from static nodes.
Normal nodes wait for the channel of the mobile device and route messages towards it, so that the communications take place in proximity (directly or at most with a limited multi-hop traversal). As a result, ordinary nodes can save energy because path length, contention and forwarding overheads are reduced as well. In addition, the mobile device can visit the network in order to extend more evenly the energy consumption due to communications. When the cost of mobilizing sensor nodes is excessive, the normal approach is to attach sensor nodes to entities that will be roaming in the sensing field anyway, such as buses or animals.

Mobility-based schemes can be classified as mobile-sink and mobile-relay schemes, depending on the type of the mobile entity. It is significant to point out here that when considering mobile schemes, an important issue is the type of control the sensor-network designer has on the mobility of nodes. A complete discussion on this point is presented by Jun et al (2005). Mobile nodes can be separated into two broad categories: they can be specifically designed as part of the network infrastructure or they can be part of the environment. When they are part of the infrastructure, their mobility can be fully controlled and are, in general, robotized.

When mobile nodes are part of the environment they might not be controllable. If they follow a strict schedule, then they have a fully predictable mobility. If not, they may have a random behavior so that no reliable assumption can be made on their mobility. In such a case, mobility patterns can be learned based on successive observations and estimated with some accuracy.

Mobility of sensor nodes is actually feasible and it can be accomplished in different ways as studied by Akyildiz et al (2004). For example, sensors can be equipped with mobilizers for changing their location. As mobilizers are generally quite expensive from the energy
consumption standpoint, adding mobility to sensor nodes may not be convenient. In fact, the resulting energy consumption may be greater than the energy gain due to the mobility itself.

So, instead of making each sensor node mobile, mobility can be limited to special nodes which are less energy constrained than the ordinary ones. In this case, mobility is strictly tied to the heterogeneity of sensor nodes. On the other hand, instead of providing mobilizers, sensors can be placed on elements which are mobile on their own (e.g. animals, cars and so on).

There are two different options in this case. First, all the sensors are put onto mobile elements, so that all nodes in the network are mobile. Alternatively, only a limited number of special nodes can be placed on mobile elements, while the other sensors are stationary. Anyway, in both cases there is no additional energy consumption overhead due to mobility, but the mobility pattern of mobile elements has to be taken into account during the network design phase (more details are provided below). As mentioned in the previous section, mobility-based energy conservation schemes can be classified depending on the nature of the mobile element, i.e. a Mobile Sink (MS) or a Mobile Relay (MR).

2.5.1 Mobile-Sink-based approaches

Many approaches proposed in the literature about sensor networks with mobile sinks (MSs) rely on a Linear Programming (LP) formulation which is exploited in order to optimize parameters such as network lifetime and so on. For example, Wang et al (2005) proposed a model consisting of an MS which can move to a limited number of locations (sink sites) to visit a given sensor and communicate with it (sensors are supposed to be arranged in a square grid within the sensing area). During visits to nodes, the sink stays at the node location for a period of time. Nodes not in the coverage area of the
sink can send messages along multi-hop paths ending at the MS and obtained using the shortest path routing.

2.5.2 Mobile-Relay-based approaches

In mobile wireless sensor networks (WSN) coverage can be enhanced by moving the sensors so that a better arrangement is achieved. However, movement is a high energy consumption task discussed by Aziz et al (2011). The Mobile Relay (MR) model for data collection in multi-hop sensor networks has already been explored in the context of opportunistic networks. One of the most well-known approaches is given by the message ferrying scheme as discussed by Zhao et al (2004). Message ferries are special mobile nodes which are introduced into a sparse mobile ad hoc network to offer the service of message relaying. Message ferries move around in the network area and collect data from source nodes. They carry stored data and forward them towards the destination node. Thus, message ferries can be seen as a moving communication infrastructure which accommodates data transfer in sparse wireless networks.

In fact, changing the trajectory of the MR is not always possible in the case of sensor networks because sensors may be deployed in places with obstacles, on rough terrain or generally where unmanned vehicles can move only in certain directions. Energy saving is addressed in such a way that a large number of nodes is visited by the MR and it can thus transmit data over a single hop connection using a short range radio. The other nodes which are not in proximity of the path followed by MR send their data over a multi-hop path. This is however shorter and thus cheaper, with respect to the path established towards a fixed sink node in a classical dense wireless sensor network.
To manage this kind of data collection, nodes self-organize into clusters where cluster heads are the nodes. These nodes are nearer to the path of the MR whereas the other nodes of the cluster send their data to the cluster head for storage until the next visit of the MR. Data from the sensor nodes of the cluster travel towards the cluster heads according to the directed diffusion protocol. Then the election of the cluster heads is kept after the first traversal of the MR.

During this traversal, the MR does not collect any data. Transmissions from cluster heads to the MR occur only when the MR is in proximity so as not to waste energy in useless transmissions. As the trajectory of the MR is assumed to be fixed, it can be controlled only in time. The MR can move at a constant speed worked out, for example, depending on the buffer constraints of the cluster heads. Each cluster head is thus visited before its buffer runs out of space. However, a better performance is experienced when the MR alternates between two states: moving at a certain constant speed or stopping. So MR moves fast in places with no or only a few sensors and stops in proximity of cluster heads where sensor deployment is denser. The determination of places where sensor deployment is denser (congested regions) is done at each traversal of the MR.

2.6 SUMMARY

In this chapter, the main approaches to energy conservation in wireless sensor networks have been surveyed. Special attention has been devoted to a systematic and comprehensive classification of the solutions proposed in the literature. All the same, the discussion of the topics that have received wide interest in the past is not limited. It also stresses the importance of different approaches such as data-driven and mobility-based schemes. It is worth noting that the considered approaches need not be construed as alternatives. They should rather be exploited together.
Later final observations about the different approaches to energy management are studied. As far as traditional techniques to energy saving are concerned, an important aspect which has to be investigated more deeply is the integration of the different approaches into a single off-the-shelf workable solution. This involves characterizing the interactions between different protocols and exploiting cross-layer interactions.

This chapter presented an overview of the present techniques for addressing energy conservation. Next, each approach along with its shortcomings has been discussed. The need for the proposed schemes has also been justified.