The Research Monograph Series in Computing, Electrical & Communication Networks
Vol. 1 (01), May 2023, pp. 15–40

# CHAPTER 2

# STATE OF THE ART

## 2.1 SENSOR ARCHITECTURAL DESIGN

Networking unattended sensor nodes are expected to have significant impact on the efficiency of many military and civil applications such as combat field surveillance, security and disaster management. These systems process data gathered from multiple sensors to monitor events in an area of interest. Sensor nodes are constrained in energy supply and bandwidth. Such constraints combined with a typical deployment of large number of sensor nodes have posed many challenges to the design and management of sensor networks. These challenges necessitate energy-awareness at all layers of networking protocol stack. The issues related to physical and link layers are generally common for all kind of sensor applications, therefore the research on these areas has been focused on system-level power awareness. At the network layer, the main aim is to find ways for energy efficient route setup and reliable relaying of data from the sensor nodes to the sink so that the lifetime of the network is maximized.

First of all, it is not possible to build a global addressing scheme for the deployment of sheer number of sensor nodes. Therefore a classical IP-based protocol cannot be applied to sensor networks.

Second, in contrary to typical communication networks almost all applications of sensor networks (Ian F. Akyildiz et al., 2002) require the flow of sensed data from multiple regions (sources) to a particular sink. Third, generated data traffic has significant redundancy in it since multiple sensors may generate same data within the vicinity of a phenomenon. Such redundancy needs to be exploited by the routing protocols to improve energy and bandwidth utilization. Fourth, sensor nodes are tightly constrained in terms of transmission power, on-board energy, processing capacity and storage and thus require careful resource management. Due to such differences, many new algorithms have been proposed for the problem of routing data in sensor networks. These routing mechanisms have considered the characteristics of sensor nodes along with the application and architecture requirements. Almost all of the routing protocols can be classified as data-centric, hierarchical or location-based although there are few distinct ones based on network flow or QoS awareness. Data-centric protocols are query-based and depend on the naming of desired data, which helps in eliminating many redundant transmissions. Hierarchical protocols aim at clustering the nodes so that cluster heads can do some aggregation and reduction of data in order to save energy. Location-based protocols utilize the position information to relay the data to the desired regions rather than the whole network. The last category includes routing approaches that are based on general network-flow modeling and protocols that strive for meeting some QOS requirements along with the routing function (Jonathan L Bredin et al., 2010).

There are three main components in a sensor network. They are the sensor nodes, sink and monitored events and mentioned below

1) An advent subsystem which sense the environment.
2) Computational logic which converts the sensed raw into computational data.
3) Message exchange protocol.

The performance of a routing protocol is closely related to the architectural model and hence whole architectural design is fully based on

- Network Dynamics
- Node Deployment
- Energy Considerations
- Data Delivery Models
- Node Capabilities
- Data Aggregation/Fusion

The design of routing protocols in WSNs is influenced by many challenging factors. These factors must be overcome before efficient communication can be achieved in WSNs. In the following, we summarize some of the routing challenges and design issues that affect routing process in WSNs.

**Node deployment:** Node deployment in WSNs is application dependent and affects the performance of the routing protocol. The deployment can be either deterministic or randomized. In deterministic deployment, the sensors are manually placed and data is routed through pre-determined paths. However, in random node deployment, the sensor nodes are scattered randomly creating an infrastructure in an ad hoc manner. If the resultant distribution of nodes is not uniform, optimal clustering becomes necessary to allow connectivity and enable energy efficient network operation. Inter-sensor communication is normally within short transmission ranges due to energy and bandwidth limitations. Therefore, it is most likely that a route will consist of multiple wireless hops.

**Energy consumption without losing accuracy:** sensor nodes can use up their limited supply of energy performing computations and transmitting information in a wireless environment. As such, energy conserving forms of communication and computation are essential. Sensor node lifetime shows a strong dependence on the battery lifetime (W. Heinzelman, et al, 2000). In a multichip WSN, each node plays a dual role as data sender and data router. The malfunctioning of some sensor nodes due to power failure can cause significant topological changes and might require rerouting of packets and reorganization of the network.

**Data Reporting Model:** Data sensing and reporting in WSNs is dependent on the application and the time criticality of the data reporting. Data reporting can be categorized as either time-driven (Continuous), event-driven, query-driven, and hybrid (Y. Yao and J. Gehrke, 2002). Data Reporting Model: Data sensing and reporting in WSNs is dependent on the application and the time criticality of the data reporting. Data reporting can be categorized as either time-driven (continuous), event-driven, query-driven, and hybrid. The time-driven delivery model is suitable for applications that require periodic data monitoring. As such, sensor nodes will periodically switch on their sensors and transmitters, sense the environment and transmit the data of interest at constant periodic time intervals. In event-driven and query-driven models, sensor nodes react immediately to sudden and drastic changes in the value of a sensed attribute due to the occurrence of a certain event or a query is generated by the BS. As such, these are well suited for time critical applications. A combination of the previous models is also possible. The routing protocol is highly influenced by the data reporting model with regard to energy consumption and route stability.

**Node/Link Heterogeneity:** In many studies, all sensor nodes were assumed to be homogeneous, i.e., having equal capacity in terms of computation, communication, and power. However, depending on the application a sensor node can have different role or capability. The existence of heterogeneous set of sensors raises many technical issues related to data routing. For example, some applications might require a diverse mixture of sensors for monitoring temperature, pressure and humidity of the surrounding environment, detecting motion via acoustic signatures, and capturing the image or video tracking of moving objects. These special sensors can be either deployed independently or the different functionalities can be included in the same sensor nodes. Even data reading and reporting can be generated from these sensors at different rates, subject to diverse quality of service constraints, and can follow multiple data reporting models. For example, hierarchical protocols designate a cluster head node different from the normal sensors. These cluster heads can be chosen from the deployed sensors or can be more powerful than other sensor nodes in terms of energy, bandwidth, and memory. Hence, the burden of transmission to the BS is handled by the set of cluster-heads.

**Fault Tolerance:** Some sensor nodes may fail or be blocked due to lack of power, physical damage, or environmental interference. The failure of sensor nodes should not affect the overall task of the sensor network. If many nodes fail, MAC and routing protocols must accommodate formation of new links and routes to the data collection base stations. This may require actively adjusting transmit powers and signaling rates on the existing links to reduce energy consumption, or rerouting packets through regions of the network where more energy is available. Therefore, multiple levels of redundancy may be needed in a fault-tolerant sensor network.

**Scalability:** The number of sensor nodes deployed in the sensing area may be in the order of hundreds or thousands, or more. Any routing scheme must be able to work with this huge number of sensor nodes. In addition, sensor network routing protocols should be scalable enough to respond to events in the environment. Until an event occurs, most of the sensors can remain in the sleep state, with data from the few remaining sensors providing a coarse quality.

**Network Dynamics:** Most of the network architectures assume that sensor nodes are stationary. However, mobility of both BS's and sensor nodes is sometimes necessary in many applications (H. Lou, et al, (2002)). Routing messages from or to moving nodes is more challenging since route stability becomes an important issue, in addition to energy, bandwidth etc. Moreover, the sensed phenomenon can be either dynamic or static depending on the application, e.g., it is dynamic in a target detection/tracking application, while it is static in forest monitoring for early fire prevention. Monitoring static events allows the network to work in a reactive mode, simply generating traffic when reporting. Dynamic events in most applications require periodic reporting and consequently generate significant traffic to be routed to the BS.

**Transmission Media:** In a multi-hop sensor network, communicating nodes are linked by a wireless medium. The traditional problems associated with a wireless channel (e.g., fading, high error rate) may also affect the operation of the sensor network. In general, the required bandwidth of sensor data will be low, on the order of 1–100 kb/s. Related to the transmission media is the design of medium access control (MAC). One approach of MAC design for sensor networks is to use TDMA based protocols that conserve more energy compared to contention based protocols like CSMA (e.g., IEEE 802.11). Bluetooth technology can also be used.

**Connectivity:** High node density in sensor networks precludes them from being completely isolated from each other. Therefore, sensor nodes are expected to be highly connected. This, however, may not prevent the network topology from being variable and the network size from being shrinking due to sensor node failures. In addition, connectivity depends on the, possibly random, distribution of nodes.

**Coverage:** In WSNs, each sensor node obtains a certain view of the environment. A given sensor's view of the environment is limited both in range and in accuracy; it can only cover a limited physical area of the environment. Hence, area coverage is also an important design parameter in WSNs.

**Data Aggregation:** Since sensor nodes may generate significant redundant data, similar packets from multiple nodes can be aggregated so that the number of transmissions is reduced. Data aggregation is the combination of data from different sources according to a certain aggregation function, e.g., duplicate suppression, minima, maxima and average. This technique has been used to achieve energy efficiency and data transfer optimization in a number of routing protocols. Signal processing methods can also be used for data aggregation. In this case, it is referred to as data fusion where a node is capable of producing a more accurate output signal by using some techniques such as beam forming to combine the incoming signals and reducing the noise in these signals.

**Quality of Service:** In some applications, data should be delivered within a certain period of time from the moment it is sensed, otherwise the data will be useless. Therefore bounded latency for data delivery is another condition for time-constrained applications. However, in many applications, conservation of energy, which is directly related to network lifetime, is considered relatively more important than the quality of data sent. As the energy gets depleted, the network may be required to reduce the quality of the results in order to reduce the energy dissipation in the nodes and hence lengthen the total network lifetime. Hence, energy-aware routing protocols are required to capture this requirement.

Routing in sensor networks is very challenging due to several characteristics that distinguish from the contemporary communication and wireless ad-hoc networks. Some of the optimal routing protocols in the Table 2.1 are discussed as follows:

## 2.2 DATA-CENTRIC PROTOCOLS

Hiker and A.Willig (2005) has mentioned that data-centric routing, the sink sends queries to certain regions and waits for data from the sensors located in the selected regions. Since data is being requested through queries, attribute based naming is necessary to specify the properties of data.

### 2.2.1 SPIN

SPIN is the first data-centric protocol, which considers data negotiation between nodes in order to eliminate redundant data. The idea behind SPIN is to name the data using high level descriptors or meta-data. Before transmission, meta-data are exchanged among sensors via data advertisement mechanism, which is the key feature of SPIN. Each node upon receiving new data, advertises it to its neighbours and interested neighbours, i.e. those who do not have the data, retrieve the data by sending a request message. SPIN's meta-data negotiation solves the classic problems of flooding such as redundant information passing, overlapping of sensing areas and resource blindness thus, achieving a lot of energy efficiency. There is no standard meta-data format and it is assumed to be application specific, e.g. using an application level framing. There are three messages defined in SPIN to exchange data between nodes. These are: ADV message to allow a sensor to advertise a particular meta-data, REQ message to request the specific data and DATA message that carry the actual data.

SPIN (Liu, et al., 2008) gives a factor of 3.5 less than flooding in terms of energy dissipation and meta-data negotiation almost halves

**Table 2.1** State of the Art – List of Routing protocols in WSN.

| **Data centric** | **SPIN** | **SPIN-PP** |
|---|---|---|
| | | SPIN-EC |
| | | SPIN-BC |
| | | SPIN-RL |
| | Direct Diffusion | |
| | Energy ware | |
| | Reliable Energy Aware Routing | |
| | Rumor | |
| | MCFA | |
| | Link Quality Estimation Based | |
| | Gradient Based | |
| | Information-driven | |
| | Acquire | |
| Hierarchical | LEACH | |
| | EWC | |
| | PEGASIS | |
| | TEEN/APTEEN | |
| | Energy-aware cluster-based | |
| | Self-organized | |
| | Minimum energy communication network | |
| | Small minimum energy communication network | |
| Location-based | Geographic Adaptive Fidelity | |
| | Energy Aware Greedy Routing | |
| | Geographic and Energy Aware | |
| QoS Aware | SPEED | |
| | MMSPEED | |
| | Sequential Assignment | |
| | Real-Time Power-Aware | |
| | DCEERP | |
| | Energy Efficient with Delay Guaranties | |

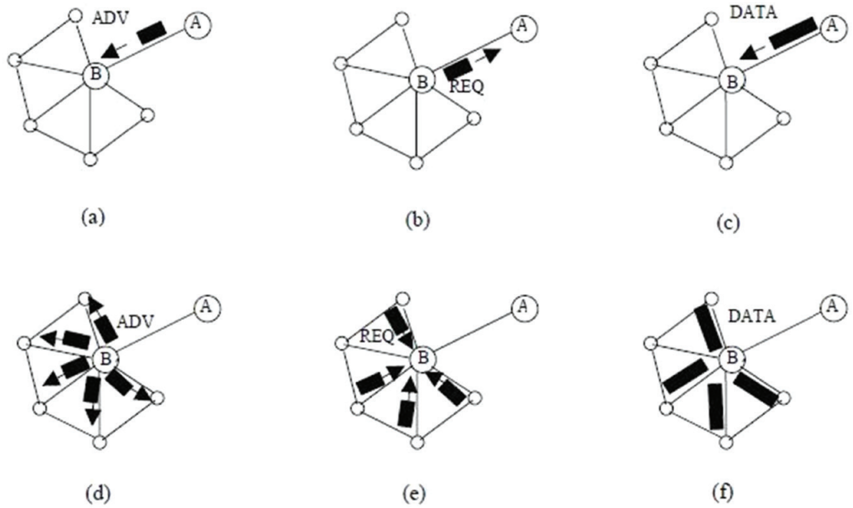


**Figure 2.1.** SPIN Protocol.

the redundant data. However, SPIN's data advertisement mechanism cannot guarantee the delivery of data. For instance, if the nodes that are interested in the data are far away from the source node and the nodes between source and destination are not interested in that data, such data will not be delivered to the destination. Figure 2.1 shows that the Node A starts by advertising its data to node B (a). Node B responds by sending a request to node A (b). After receiving the requested data (c), node B then sends out advertisements to its neighbour's (d), who in turn send requests back to B (e–f).

SPIN is one of the most important data centric protocols in WSN which operates in two ways. Firstly, to complete the task effectively and to conserve energy in the data processing, then it establishes the request and response throughout the base about the data which was already in the sensor nodes. Secondly, each node should adapt to the network about it changes and resource utilization. Metadata is collected for each sensor node and these metadata's are used to describe or to reference the sensor data. Meta data does not have a specific format and mostly the Meta data format is application orientation. There are mainly three adverse flag of SPIN protocol,

| | | |
|---|---|---|
| ADV | - | advertisement flag which broadcast the meta |
| data REQ | - | request flag to send particular request |
| DATA | - | message flag that hold on the data values (real entities not Meta data) |

**SPIN-PP**: This protocol is designed mainly for peer to peer communication, using this protocol two sensor node or node to base station can communicate to each other without any interference.

**SPIN-EC**: This protocol uses 3 way handshaking with heuristic conservation. A node can make communication only with the active participation of the active node with certain threshold. Generally node with low energy level will sends REQ message along with data message.

**SPIN-BC**: This protocol is used for broadcasting within the network with shared channel. This protocol initially broadcast all the message within a certain range. This protocol works on basis of the three conditions mentioned below

- If a node receives an ADV advertisement message, it doesn't respond with REQ message quickly, it has to wait for certain period of time.
- If a node, except the broadcasting node receives REQ flag, the request sent from the particular node gets cancelled.
- If an advertising node with ADV flag receives the REQ flag then the node sends only the data message

**SPIN-RL**: This protocol is as same as SPIN-BC, here the node always track the ADV message, if the message is not received at the receiver end, then the node again sends the REQ message to gain the access with limited frequency along with time trial.

### 2.2.2 Flooding

Flooding is a classical mechanism to relay data in sensor networks without the need for any routing algorithms and topology maintenance. In flooding, each sensor receiving a data packet broadcasts it to all of its neighbours and this process continues until the packet arrives at the destination or the maximum number of hops for the packet is reached which shown in the Figure 2.2. Lieckfeldt, D. et al. (2008), claimed the main drawback of flooding protocol is implosion caused by duplicated messages sent to same node, overlap when two nodes sensing the same region send similar packets to the same neighbour and resource blindness by consuming large amount of energy without consideration for the energy constraints.

### 2.2.3 Directed Diffusion

The idea aims at diffusing data through sensor nodes by using a naming scheme for the data. The main reason behind using such a scheme

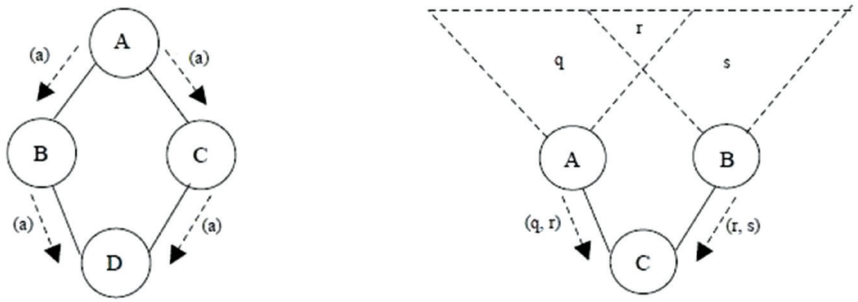


**Figure 2.2.** Flooding along with impulsion and overlapping problem.

is to get rid of unnecessary operations of network layer routing in order to save energy. Direct Diffusion (Marco Źũniga Z. and Bhaskar Krishnamachari) suggests the use of attribute-value pairs for the data and queries the sensors in an on demand basis by using those pairs. In order to create a query, an interest is defined using a list of attribute-value pairs such as name of objects, interval, duration, geographical area, etc. In Figure 2.3 the interest is broadcast by a sink through its neighbour's. Each node receiving the interest can do caching for later use. The nodes also have the ability to do in-network data aggregation, which is modeled as a minimum Steiner tree problem.

Directed Diffusion differs from SPIN in terms of the on demand data querying mechanism. In Directed Diffusion the sink queries the sensor nodes if a specific data is available by flooding some tasks (Marron, P.J. and Minder, D., 2006). In SPIN, sensors advertise the availability of data allowing interested nodes to query that data.

Directed Diffusion has many advantages. Since it is data centric, all communication is neighbour-to-neighbour with no need for a node addressing mechanism. Each node can do aggregation and caching, in addition to sensing. Caching is a big advantage in terms of energy efficiency

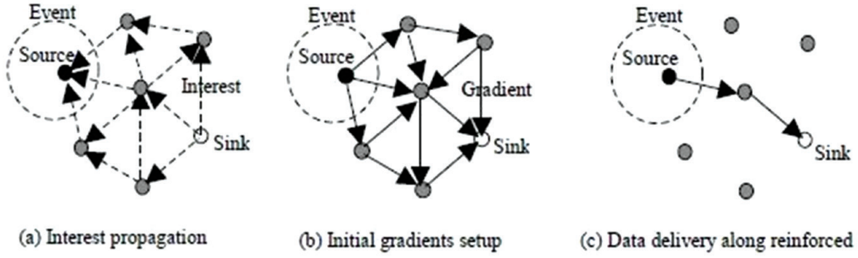


**Figure 2.3.** Directed Diffusion routing mechanism.

and delay. In addition, Direct Diffusion is highly energy efficient since it is on demand and there is no need for maintaining global network topology.

In this protocol, the data generated from each sensor nodes are diffused within the sensor nodes with a naming scheme to avoid unnecessary routing happening in the sensor network at network layer to increase power efficiency. In this protocol each attribute values are notified and are requested to use the particular data pair and query the data on basis of on demand principle. Here when ever demand arises the query is submitted by requesting the data pair, the query is created based on the defined interest using the attribute values of each data pair. Figure 2.4 shows about the broadcasting medium by the sink nodes along with caching and gradient field. Gradient field are used as reply link to the neighbour node through which it receives the interest of the query. The cached interests are compared with the real time query interest to characterize the data rate, processing, packet duration etc. When the node failure is achieved between the sink and sources, the alternate path is traced immediately to prove the aptness of the protocol.

However, directed diffusion cannot be applied to all sensor network applications since it is based on a query-driven data delivery model (Park, S. J., et al., 2004). The applications that require continuous data delivery to the sink will not work efficiently with a query-driven on demand data model. Therefore, Directed Diffusion is not a good choice as a routing protocol for the applications such as environmental monitoring. In addition, the naming schemes used in Directed Diffusion are application dependent and each time should be defined a priori. Moreover, the matching process for data and queries might require some extra overhead at the sensors.

## 2.2.4 Energy-Aware Routing

The main key idea is to use a set of sub-optimal paths occasionally to increase the lifetime of the network. These paths are chosen by means of a probability function (Patwari, N. et al., 2003), which depends on the

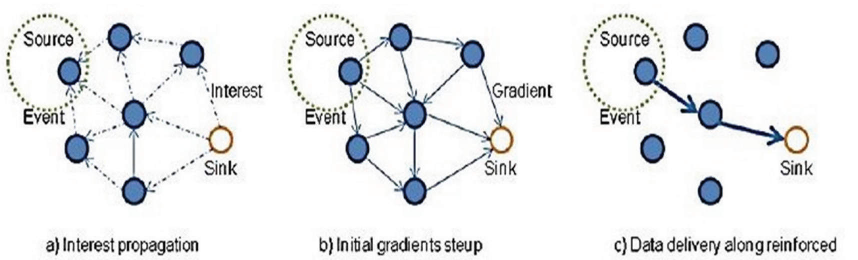


**Figure 2.4.** Directed Diffusion data delivery mechanism.

energy consumption of each path. Network survivability is the main metric that the approach is concerned with. The approach argues that using the minimum energy path (Rahman, M. A., et al., 2008) all the time will deplete the energy of nodes on that path. Instead, one of the multiple paths is used with a certain probability so that the whole network lifetime increases. The protocol assumes that each node is addressable through a class-based addressing which includes the location and types of the nodes (Savvides, C.-C. Han and Strivastava, M. B., 2001). There are 3 phases in the protocol:

- Setup phase: Localized flooding occurs to find the routes and create the routing tables. While doing this; the total energy cost is calculated in each node.
- Data Communication Phase: Each node forwards the packet by randomly choosing a node from its forwarding table using the probabilities.
- Route maintenance phase: Localized flooding is performed infrequently to keep all the paths alive.

Energy aware routing approach is similar to Directed Diffusion in the way potential paths from data sources to the sink are discovered (Schurgers, C., et al., 2002). In Directed Diffusion, data is sent through multiple paths, one of them being reinforced to send at higher rates. On the other hand, Shah et al. select a single path randomly from the multiple alternatives in order to save energy. Therefore, when compared to Directed Diffusion, it provides an overall improvement of 21.5% energy saving and a 44% increase in network lifetime.

### 2.2.5 Rumor Routing

Rumor Routing is another variation of Directed Diffusion and is mainly intended for contexts in which geographic routing criteria are not applicable. Generally Directed Diffusion floods the query to the entire network when there is no geographic criterion to diffuse tasks. However, in some cases there is only a little amount of data requested from the nodes and thus the use of flooding is unnecessary (Shinji Motegi, et al., 2005). An alternative approach is to flood the events if number of events is small and number of queries is large. Rumor routing is between event flooding and query flooding. The idea is to route the queries to the nodes that have observed a particular event rather than flooding the entire network to retrieve information about the occurring events.

In order to flood events through the network, the rumor routing algorithm employs long lived packets, called agents. When a node detects an event, it adds such event to its local table and generates an agent (Shneidman, J., et al., 2014). Agents travel the network in order to propagate information about local events to distant nodes. When a node generates a query for an event, the nodes that know the route, can respond to the query by referring its event table. Hence, the cost of flooding the whole network is avoided. Rumor routing maintains only one path between source and destination as opposed to Directed Diffusion where data can be sent through multiple paths at low rates.

### 2.2.6 Gradient-Based Routing

The idea is to keep the number of hops when the interest is diffused through the network. Hence, each node can discover the minimum number of hops to the sink, which is called height of the node. The difference between a node's height and that of its neighbour is considered the gradient on that link. A packet is forwarded on a link with the largest gradient.

### 2.2.7 Constrained Anisotropic Diffusion Routing (CADR)

The idea is to query sensors and route data in a network in order to maximize the information gain, while minimizing the latency and bandwidth. This is achieved by activating only the sensors that are close to a particular event and dynamically adjusting data routes (Sohraby, K et al., 2007). The major difference from Directed Diffusion is the consideration of information gain in addition to the communication cost. In CADR, each node evaluates an information/cost objective and routes data based on the local information/cost gradient and end-user requirements. The information utility measure is modeled using standard estimation theory.

### 2.2.8 Information-Driven Sensor Querying (IDSQ)

IDSQ is based on a protocol in which the querying node can determine which node can provide the most useful information while balancing the energy cost. While IDSQ provides away of selecting the optimal order of sensors for maximum incremental information gain, it does not specifically define how the query and the information are routed between sensors and the sink. Therefore, IDSQ (Tilak S. et al., 2002) can be seen as a complementary optimization procedure.

### 2.2.9 COUGAR Routing

The main idea is to use declarative queries in order to abstract query processing from the network layer functions such as selection of relevant sensors etc. and utilize in-network data aggregation to save energy. The abstraction is supported through a new query layer between the network and application layers.

### 2.2.10 ACtiveQUery forwarding InsensoRnEtworks (ACQUIRE)

This approach views the sensor network as a distributed database and is well-suited for complex queries which consist of several sub queries. The querying mechanism works as follows: The query is forwarded by the sink and each node receiving the query, tries to respond partially by using its pre-cached information and forward it to another sensor. If the pre-cached information is not up-to-date, the nodes gather information from its neighbour's within a look-ahead of *d* hops. Once the query is being resolved completely, it is sent back through either the reverse or shortest-path to the sink. One of the main motivations for proposing ACQUIRE is to deal with one-shot, complex queries for data where a response can be provided by many nodes. Since, the data-centric approaches such as Directed Diffusion uses flooding-based query mechanism for continuous and aggregate queries; it would not make sense to use the same mechanism for one-shot complex queries due to energy considerations (Wang, K. Sohraby et al., 2006).

## 2.3 HIERARCHICAL PROTOCOLS

The main aim of hierarchical routing is to efficiently maintain the energy consumption of sensor nodes by involving them in multi-hop communication within a particular cluster and by performing data aggregation and fusion in order to decrease the number of transmitted messages to the sink. Cluster formation is typically based on the energy reserve of sensors and sensor's proximity to the cluster head.

### 2.3.1 Low-Energy Adaptive Clustering Hierarchy (LEACH)

The idea is to form clusters of the sensor nodes based on the received signal strength and use local cluster heads as routers to the sink. This will save energy since the transmissions will only be done by such cluster heads rather than all sensor nodes. Optimal number of cluster heads is estimated to be 5% of the total number of nodes. All the data processing such as data fusion and aggregation are local to the cluster (Wang, et al., 2013). Cluster heads change randomly over time in order to balance the energy dissipation of nodes. This decision is made by the node choosing a random number between 0 and 1. The node becomes a cluster head for the current round if the number is less than the threshold.

### 2.3.2 Power-Efficient Gathering in Sensor Information Systems (PEGASIS)

The key idea is PEGASIS forms chains from sensor nodes so that each node transmits and receives from a neighbour and only one node is selected from that chain to transmit to the base station (sink).

Gathered data moves from node to node, aggregated (Wei Ye et al., 2002) and eventually sent to the base station. The chain construction is performed in a greedy way.

### 2.3.3 Hierarchical Power-Efficient Gathering in Sensor Information Systems (Hierarchical PEGASIS)

Hierarchical PEGASIS is an extension to PEGASIS, which aims at decreasing the delay incurred for packets during transmission to the base station and proposes a solution to the data gathering problem by considering energy × delay metric. In order to reduce the delay in PEGASIS, simultaneous transmissions of data messages are pursued (Wendi Rabiner Heinzelman, et al, 2000). To avoid collisions and possible signal interference among the sensors, two approaches have been investigated. The first approach incorporates signal coding, e.g. CDMA. In the second approach only spatially separated nodes are allowed to transmit at the same time. The chain-based protocol with CDMA capable nodes, constructs a chain of nodes, that forms a tree like hierarchy, and each selected node in a particular level transmits data to the node in the upper level of the hierarchy. This method ensures data transmitting in parallel and reduces the delay significantly (Xiang Yang Li, 2008).

### 2.3.4 Threshold sensitive Energy Efficient sensor Network protocol (TEEN)

TEEN is a hierarchical protocol designed to be responsive to sudden changes in the sensed attributes such as temperature. Responsiveness is important for time-critical applications, in which the network operated in a reactive mode. TEEN pursues a hierarchical approach along with the use of a data-centric mechanism. The sensor network architecture is based on a hierarchical grouping where closer nodes form clusters and this process goes on the second level (Ye, F., 2001). After the clusters are formed, the cluster head broadcasts two thresholds to the nodes. These are hard and soft thresholds for sensed attributes. Hard threshold is the minimum possible value of an attribute to trigger a sensor node to switch on its transmitter and transmit to the cluster head. Thus, the hard threshold allows the nodes to transmit only when the sensed attribute is in the range of interest, thus reducing the number of transmissions significantly. Once a node senses a value at or beyond the hard threshold, it transmits data only when the values of that attribute changes by an amount equal to or greater than the soft threshold (Younis, M. et al., 2002). As a consequence, soft threshold will further reduce the number of transmissions if there is little or no change in the value of sensed attribute. One can adjust both hard and soft threshold values in order to control the number of packet transmissions. However, TEEN is not good for applications where periodic reports are needed since the user may not get any data at all if the thresholds are not reached, until base station (sink) is reached.

### 2.3.5 Adaptive Threshold sensitive Energy Efficient sensor Network protocol (APTEEN)

APTEEN is an extension to TEEN and aims at both capturing periodic data collections and reacting to time-critical events. The architecture is same as in TEEN. When the base station forms the clusters, the cluster heads broadcast the attributes, the threshold values, and the transmission schedule to all nodes. Cluster heads also perform data aggregation in order to save energy (Younis, M. et al., 2002). APTEEN supports three different query types: historical, to analyze past data values one-time, to take a snapshot view of the network; and persistent to monitor an event for a period of time.

### 2.3.6 Energy-Aware Routing for Cluster-Based Sensor Networks

Sensors are grouped into clusters prior to network operation. The algorithm employs cluster heads, namely gateways, which are less energy constrained than sensors and assumed to know the location of sensor nodes. Gateways maintain the states of the sensors and sets up multi-hop routes for collecting sensors' data. A TDMA based MAC is used for nodes to send data to the gateway. The gateway informs each node about slots in which it should listen to other nodes' transmission and slots, which the node can use for its own transmission. The command node (sink) communicates only with the gateways.

The sensor is assumed to be capable of operating in an active mode or a low-power stand-by mode. The sensing and processing circuits can be

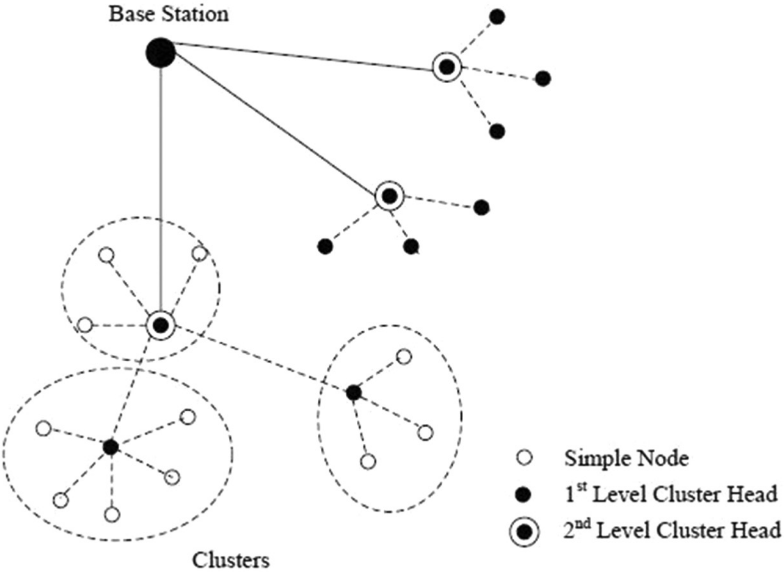


**Figure 2.5.** Routing Mechanism in TEEN and APTEEN.

powered on and off. In addition both the radio transmitter and receiver can be independently turned on and off and the transmission power can be programmed based on the required range (Chu, M., et al., 2002). The sensor nodes in a cluster can be in one of four main states: sensing only, relaying only, sensing-relaying, and inactive. In the sensing state, the node probes the environment and generates data at a constant rate. In the relaying state, the node does not sense the target but its communications circuitry is on to relay the data from other active nodes. When a node is both sensing and relaying messages from other nodes, it is considered in the sensing-relaying state. Otherwise, the node is considered inactive and can turn off its sensing and communication circuitry.

## 2.4 SELF-ORGANIZING PROTOCOL

Adam Dunkels, et al. (2004) have proposed an architecture that supports heterogeneous sensors that can be mobile or stationary. C. Schurgers et al. (2002) have discovered some sensors, which can be either stationary or mobile, probe the environment and forward the data to designated set of nodes that act as routers. Router nodes are stationary and form the backbone for communication. Collected data are forwarded through the routers to more powerful sink nodes. Each sensing node should be reachable to a router node in order to be part of the network.

A routing architecture that requires addressing of each sensor node has been addressed. Sensing nodes are identifiable through the address of the router node it is connected to. The routing architecture is hierarchical where groups of nodes are formed and merge when needed. In order to support fault tolerance, Local Markov Loops (LML) algorithm, which performs a random walk on spanning trees of a graph, is used in broadcasting. The algorithm for self organizing the router nodes and creating the routing tables consists of four phases:

- Discovery phase: Y. Thomas Hou et al. (2005) has proposed the nodes in the neighborhood of each sensor are discovered.
- Organization phase: Jonathan L Bredin et al. (2010) had mentioned groups are formed and merged by forming a hierarchy. Each node is allocated an address based on its position in the hierarchy. Routing tables of size O(log N) are created for each node. Broadcast trees that span all the nodes are constructed.
- Maintenance phase: Y. Thomas Hou, et al. (2005) has mentioned that updating of routing tables and energy levels of nodes is made in this phase. Each node informs the neighbors about its routing table and energy level. LML are used to maintain broadcast trees.
- Self-reorganization phase: In case of partition or node failures, group reorganizations are performed.

## 2.5 LOCATION-BASED PROTOCOLS

Most of the routing protocols for sensor networks require location information for sensor nodes. In most cases location information (Garcia-Alfaro, J., et al., 2010) is needed in order to calculate the distance between two particular nodes so that energy consumption can be estimated. Since there is no addressing scheme for sensor networks like IP-addresses and they are spatially deployed on a region, location information can be utilized in routing data in an energy efficient way. For instance, if the region to be sensed is known, using the location of sensors, the query can be diffused only to that particular region which will eliminate the number of transmission significantly. Some of the protocols discussed here are designed primarily for mobile ad hoc networks.

### 2.5.1 Minimum Energy Communication Network (MECN)

MECN maintains a minimum energy network for wireless networks by utilizing low power GPS. Although, the protocol assumes a mobile network, it is best applicable to sensor networks which are not mobile.

A minimum power topology for stationary nodes including a master node is found. MECN assumes a master-site as the information sink, which is always the case for sensor networks. Figure 2.6 shows the MECN identifies a relay region for every node. The relay region consists of nodes in a surrounding area where transmitting through those nodes is more energy efficient than direct transmission.

The main idea of MECN (Karim, L. and Nasser, N., 2012) is to find a sub-network, which will have less number of nodes and require less power for transmission between any two particular nodes. In this way global minimum power paths are found without considering all the nodes in the network. This is performed using a localized search for each node considering its relay region. The protocol has two phases:

1) It takes the positions of a two dimensional plane and constructs a sparse graph (enclosure graph), which consists of all the enclosures of each transmit node in the graph. This construction requires local computations in the nodes. The enclose graph contains globally optimal links in terms of energy consumption.
2) Finds optimal links on the enclosure graph. It uses distributed Belmann-Ford shortest path algorithm with power consumption as the cost metric. In case of mobility the position coordinates are updated using GPS. MECN is self-reconfiguring and thus can dynamically adapt to node's failure or the deployment of new sensors. Between two successive wake-ups of the nodes, each node can execute the first phase of the algorithm and the minimum cost links are updated by considering leaving or newly joining nodes.

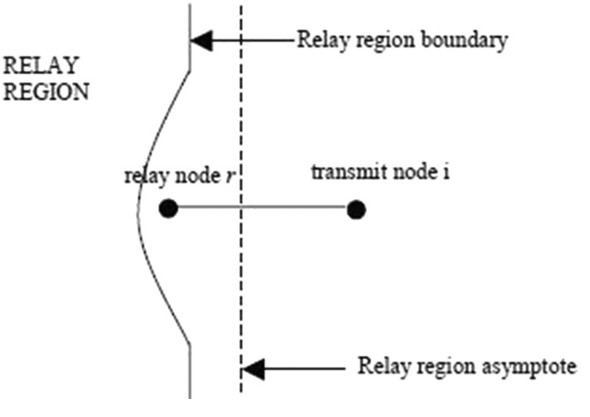


**Figure 2.6.** Relay regions in MECN and SMECN.

### 2.5.2 Small minimum energy communication network (SMECN)

SMECN is an extension to MECN. In MECN it is assumed that every node can transmit to every other node which is not possible every time. In SMECN possible obstacles between any pair of nodes are considered. However, the network is still assumed to be fully connected as in the case of MECN (Karim, L. and Nasser, N., 2012)). The sub network constructed by SMECN for minimum energy relaying is provably smaller (in terms of number of edges) than the one constructed in MECN if broadcasts are able to reach to all nodes in a circular region around the broadcaster. As a result the number of hops for transmissions will decrease. Simulation results show that SMECN uses less energy than MECN and maintenance cost of the links is less. However finding a sub-network with smaller number of edges introduces more overhead in the algorithm.

### 2.5.3 Geographic Adaptive Fidelity (GAF)

GAF is an energy-aware location-based routing algorithm designed primarily for mobile ad hoc networks (Karl, H. and Willig, A., 2005), but may be applicable to sensor networks. GAF conserves energy by turning off unnecessary nodes in the network without affecting the level of routing fidelity. It forms a virtual grid for the covered area. Each node uses its GPS-indicated location to associate itself with a point in the virtual grid. Nodes associated with the same point on the grid are considered equivalent in terms of the cost of packet routing. Such equivalence is exploited in keeping some nodes located in a particular grid area in sleeping state in order to save energy. Thus GAF can substantially increase the network lifetime as the number of nodes increases.

Nodes change states from sleeping to active in turn so that the load is balanced. There are three states defined in GAF. These states are discovery, for determining the neighbours in the grid, active reflecting participation in routing and sleep when the radio is turned off. The state transitions in GAF are depicted in Figure 2.7. Which node will sleep for how long is application dependent and the related parameters are tuned accordingly during the routing process. In order to handle the mobility, each node in the grid estimates it's leaving time of grid and sends this to its neighbour's (Karl, H. and Willig, A., 2005). The sleeping neighbour's adjust their sleeping time accordingly in order to keep the routing fidelity. Before the leaving time of the active node expires, sleeping nodes wake up and one of them becomes active. GAF is implemented both for non-mobility (GAF-basic) and mobility (GAF- mobility adaptation) of nodes.

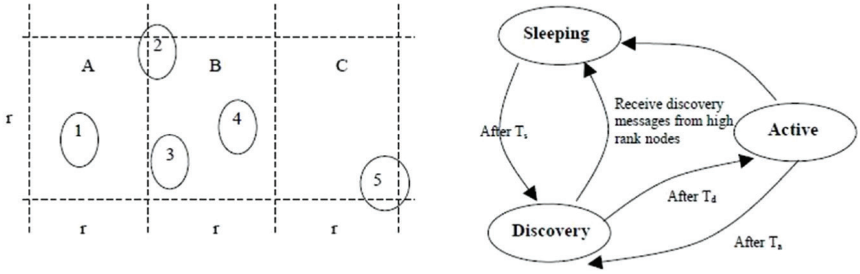


**Figure 2.7.** GAF mechanism of various states.

## 2.5.4 Geographic and Energy Aware Routing (GEAR)

GEAR uses energy aware and geographically informed neighbour selection heuristics to route a packet towards the target region. The idea is to restrict the number of interests in Directed Diffusion by only considering a certain region rather than sending the interests to the whole network. GEAR compliments Directed Diffusion in this way and thus conserves more energy. In GEAR each node keeps an estimated cost and a learning cost of reaching the destination through its neighbour's. The estimated cost is a combination of residual energy and distance to destination.

The learned cost is a refinement of the estimated cost that accounts for routing around holes in the network (Kevin, 2003). A hole occurs when a node does not have any closer neighbour to the target region than itself. If there are no holes, the estimated cost is equal to the learned cost. The learned cost is propagated one hop back every time a packet reaches the destination so that route setup for next packet will be adjusted. There are two phases in the algorithm.

1) Forwarding packets towards the target region: Upon receiving a packet, a node checks its neighbour's to see if there is one neighbour which is closer to the target region than itself. If there is more than one (Kevin, 2003), the nearest neighbour to the target region is selected as the next hop. If they are all further than the node itself, this means there is a hole. In this case, one of the neighbours is picked to forward the packet based on the learning cost function. This choice can be updated according to the convergence of the learned cost during the delivery of packets.
2) Forwarding the packets within the region: If the packet has reached the region, it can be diffused in that region by either recursive geographic forwarding or restricted flooding. Restricted flooding

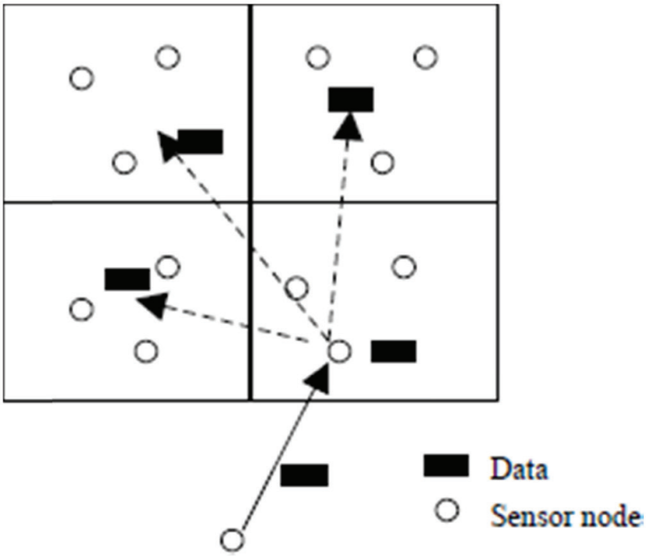


**Figure 2.8.** GEAR Routing Mechanism – recursive forwarding mechanism.

is good when the sensors are not densely deployed. In high-density networks, recursive geographic flooding is more energy efficient than restricted flooding. In that case, the region is divided into four sub regions and four copies of the packet are created. This splitting and forwarding process continues until the regions with only one node are left.

## 2.6 NETWORK FLOW PROTOCOLS

In some approaches, route setup is modeled and solved as a network flow problem. QoS-aware protocols consider end-to end delay requirements while setting up the paths in the sensor network.

### 2.6.1 Maximum Lifetime Energy Routing

The main objective of the approach is to maximize the network lifetime by carefully defining link cost as a function of node remaining energy and the required transmission energy using that link. Finding traffic distribution is a possible solution to the routing problem in sensor networks and based on that, comes the name "maximum lifetime energy routing". The solution to this problem maximizes the feasible time the network lasts.

## 2.6.2 Maximum Lifetime Data Gathering

Maximum lifetime data gathering models the data routes setup in sensor networks as the maximum lifetime data-gathering problem and presents a polynomial time algorithm. The lifetime "T" of the system is defined as the number of rounds or periodic data readings from sensors until the first sensor dies. The data-gathering schedule specifies for each round how to get and route data to the sink (Thomas Hou, Y., et al., 2005). A schedule has one tree for each round which is directed from the sink and spans all the nodes in the system. The system lifetime depends on the duration for which the schedule remains valid. The aim is to maximize the lifetime of the schedule.

## 2.6.3 Maximum Lifetime Data Aggregation (MLDA)

The algorithm considers data aggregation while setting up maximum lifetime routes. In this case, if a schedule "S" with "T" rounds is considered it induces a flow network G. The flow network with maximum lifetime subject to the energy constraints of sensor nodes is called an optimal admissible flow network (Rahman, M. A., et al., 2008). Then, a schedule is constructed by using this admissible flow network. A variant of the problem is also considered for the applications where data aggregation is not possible, i.e. steams from video sensors (Thomas Hou, Y., et al., 2005). In this case, the scenario is called Maximum Lifetime Data Routing (MLDR) and is modeled as a network flow problem with energy constraints on sensors.

## 2.6.4 Minimum Cost Forwarding

Minimum cost forwarding aims at finding the minimum cost path in a large sensor network, which will also be simple and scalable. The protocol is not really flow-based, however since data flows over the minimum cost path (Savvides, C.-C. Han and Strivastava, M. B., 2001) and the resources on the nodes are updated after each flow, we have included it in this section. The cost function for the protocol captures the effect of delay, throughput and energy consumption from any node to the sink. There are two phases in the protocol.

- First phase is a setup phase for setting the cost value in all nodes. It starts from the sink and diffuses through the network. Every node adjusts its cost value by adding the cost of the node it received the message from and the cost of the link. Such cost adjustment is not done through flooding. Instead, a back-off based algorithm is used in order to limit the number of messages exchanged. The forwarding of message is deferred for a present duration to allow the message with a minimum cost to arrive. Hence, the algorithm finds optimal cost of all nodes to the sink by using only one message at each node. Once these cost fields are set, there will be no need to keep next hop states for the nodes. This will ensure scalability.

- Second phase, the source broadcasts the data to its neighbours. The nodes receiving the broadcast message, adds its transmission cost (to sink) to the cost of the packet. Then the node checks the remaining cost in the packet. If the remaining cost of the packet is not sufficient to reach the sink, the packet is dropped. Otherwise the node forwards the packet to its neighbours. The protocol does not require any addresses and forwarding paths.

## 2.7 QoS-AWARE PROTOCOL

### 2.7.1 Sequential Assignment Routing (SAR)

SAR is the first protocol for sensor networks that includes the notion of QoS in its routing decisions. It is a table-driven multi-path approach striving to achieve energy efficiency and fault tolerance (Schurgers, C., et al., 2002). The SAR protocol creates trees rooted at one-hop neighbours of the sink by taking QoS metric, energy resource on each path and priority level of each packet into consideration. By using created trees, multiple paths from sink to sensors are formed. One of these paths is selected according to the energy resources and QoS on the path. Failure recovery is done by enforcing routing table consistency between upstream and downstream nodes on each path. Any local failure causes an automatic path restoration procedure locally. Simulation results show that SAR offers less power consumption than the minimum-energy metric algorithm, which focuses only the energy consumption of each packet without considering its priority. SAR maintains multiple paths from nodes to sink. Although, this ensures fault-tolerance and easy recovery, the protocol suffers from the overhead of maintaining the tables and states at each sensor node especially when the number of nodes is shuge.

### 2.7.2 Energy-Aware QoS Routing Protocol

Energy-Aware QoS Routing Protocol (Thomas Hou, Y., et al., 2005) main idea is to find a least cost and energy efficient path that meets certain end-to end delay during the connection. The link cost used is a function that captures the nodes' energy reserve, transmission energy, error rate and other communication parameters. In order to support both best effort and real time traffic at the same time, a class-based queuing model is employed (Wei Ye, John Heidemann and Deborah Estrin, 2002). The queuing model allows service sharing for real-time and non-real-time traffic.

The bandwidth ratio *r*, is defined as an initial value set by the gateway and represents the amount of bandwidth to be dedicated both to the real-time and non-real-time traffic on a particular outgoing link in case of a congestion. As a consequence, the throughput for normal data does not diminish by properly adjusting such "*r*" value. The queuing model is depicted in Figure 2.9. The protocol finds a list of least cost paths by using an extended version of Dijkstra's algorithm and picks a path from that list which meets the end-to-end delay requirement Protocol (Thomas Hou, Y., et al., 2005).

### 2.7.3 Speed

The protocol requires each node to maintain information about its neighbours and uses geographic forwarding to find the paths. In addition, SPEED strive to ensure a certain speed for each packet in the network so

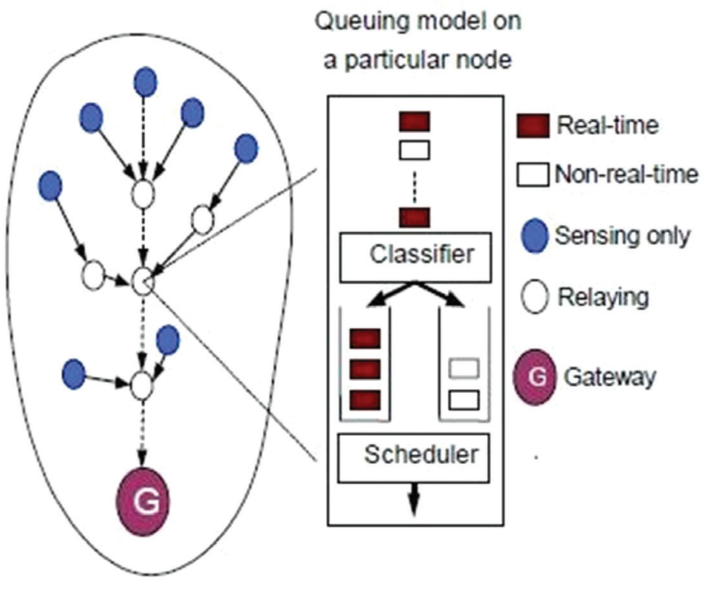


**Figure 2.9.** Energy-Aware QoS Routing Protocol – Queuing model in sensor node.

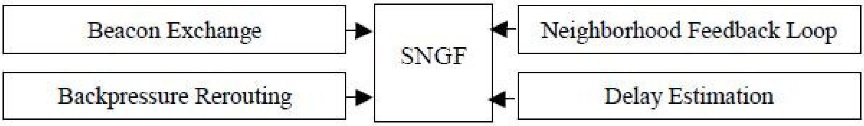


**Figure 2.10.** Routing in SPEED – SNFG module.

that each application can estimate the end-to-end delay for the packets by dividing the distance to the sink by the speed of the packet before making the admission decision. Moreover, SPEED (Wei Ye, John Heidemann et al., 2002) can provide congestion avoidance when the network is congested. The routing module in SPEED is called SNFG.

Stateless Geographic Non-Deterministic forwarding (SNFG) and works with four other modules at the network layer the beacon exchange mechanism collects information about the nodes and their location. Delay estimation at each node is basically made by calculating the elapsed time when an ACK is received from a neighbour as a response to a transmitted data packet. By looking at the delay values, SNGF selects the node which meets the speed requirement. If such a node cannot be found, the relay ratio of the node is checked is depicted in Figure 2.10. The Neighbourhood Feedback Loop module is responsible for providing the relay ratio which is calculated by looking at the miss ratios of the neighbours of a node (the nodes which could not provide the desired speed) and is fed to the SNGF module Protocol (Thomas Hou, Y, et al., 2005). If the relay ratio is less than a randomly generated number between 0 and 1, the packet is dropped. Finally the backpressure-rerouting module is used to prevent voids when a node fails to find a next hop node and to eliminate congestion by sending messages back to the source nodes so that they will try new routes.