Chapter 2

Literature Survey

This chapter deals with one of the applications using sensors as in agriculture. Details of one of the ongoing projects COMMON Sense Net (Community Oriented Monitoring and Management of Natural resources via Sensor Network, in short CSN) at Chennakeshavapura, Pavagad, Tumkur District, Karnataka are studied. Further, other concepts, like routing, aggregation, topology formation energy consumption required for the proposed work is studied.

2.1 Application Background

An optimized water management in agriculture is needed. Manual collection of soil moisture data at different depths for duration of six months is a cumbersome and unreliable process. Optimized water management means better effects on the soil’s and plants’ condition. Monitoring the level and quality of groundwater is also critical. Sensor networks can provide this added-value information. Hence a system is designed and implemented so that it is affordable, reliable, scalable and replicable in agricultural management focused on water.

Deploying automated sensors in well-identified points and let them collect data over the period could simplify the process and provide data specific to different field conditions. Once the sensor network is deployed, the data is gathered repetitively, saved into a database uploaded regularly and information utilized as and when necessary by crop modeling specialists. Marginal farmers can benefit from the technology of deficit irrigation, an agricultural water management system in which the water needs of the crop, during the growing period can only be met
partially by a combination of soil water, rainfall and irrigation. Real-time alerts are
given whenever the measured soil-moisture reaches a threshold in the benchmark
values. These alerts are automated, but farmers have to be notified by the system
operator. Once the alert is given, the farmer should be able to look at weather
forecast and know based on historical climatic data for the region, what is the
probability of rain in the near future. Based on the same type of request as above,
the system gives an estimation of the minimum amount of irrigation water needed
according to the benchmarks. A network is created to centralize the collection of
information and the farmer could access the reading from the access point as and
when desired. Figure 2.1 shows the overview of the CSN project.

Fig. 2.1: An Overview of CSN Project
[Courtesy: http://commonsense.epfl.ch/Resources/Docs/commonSensePoster.pdf]
2.2 Technical Details

Sensors need external power to run their electronic circuitry. To drive the network, power can be obtained from the battery, power grid lines or solar panel. Fields are normally located outside the village limits and it is difficult to get a continuous source of supply to these sensors. Solar panels get blocked by large sized leaves or the equipment could be more attractive to be stolen. Hence battery cell is the next choice.

Data must be collected periodically, to assess continuously the effectiveness of water conservation measures, or to run predictive scenarios about the soil moisture level. Data collection should be such that it should be in large amount, while not compromising on the lifetime of the network, so that it remains operational throughout a full season at the minimum. A periodic data collection model, with a variable rate of acquisition depending on the data variations is created. It is not necessary to collect soil moisture data at more than one sample per hour, or even per day, when there is no rainfall. However, when water is brought in either by precipitation or irrigation, a finer resolution might be desirable.

Table 2.1 shows the comparison of Tinynode (Tinynode 2011) from Shockfish with Mica (Mica 2009) motes from Crossbow. Sensor nodes from Tinynode is used in the current study as the range of transmission is larger than Mica and performs better in sleep states as shown in Table 2.2 (Panchard et al. 2009). The author proposed the usage of Tinynode motes as against the Mica motes due to the large transmission range and relatively low current consumption.

Table 2.1 Comparison of Transmission Range of Sensors

<table>
<thead>
<tr>
<th>Motes</th>
<th>Micaz</th>
<th>Mica2</th>
<th>Tinynode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transmission range in meters using an antenna of height 1 meter</td>
<td>100</td>
<td>152</td>
<td>200</td>
</tr>
</tbody>
</table>
# Table 2.2 Comparison of Energy Consumption

<table>
<thead>
<tr>
<th>Current consumption</th>
<th>Sleep</th>
<th>Transit</th>
<th>Receive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tinynode</td>
<td>1 μA</td>
<td>33 mA</td>
<td>14 mA</td>
</tr>
<tr>
<td>Mica2</td>
<td>&lt;15 μA</td>
<td>27 mA</td>
<td>10 mA</td>
</tr>
</tbody>
</table>

## 2.3 Crop and WSN

Crop field monitoring is a useful class of sensor network applications with enormous potential benefits for the farmers and society as a whole. Work related to agriculture is carried out using Fleck mote with their placement in the form of a grid by (Wark et al. 2007). (Baggio 2005) showed the usage of Lofar mote in the agricultural purpose. (Yu et al. 2013) suggested a hybrid system where the network is established under the soil and a mobile sink is used to gather the information. (Camilli et al. 2007) proposed a wireless sensor network capable of providing on-the-fly field parameter estimates for end users using local data and attaining the required precision. A cut off values for sensed information is set up and it is assigned a value 0 below cutoff and 1 above cut off. A border is created around an area of correlated data and data is sent on query request. (Javier et al. 2011) proposed data monitoring system for crops which are distributed. Clusters are created for each area of the crop. These clusters send the data to a common access point. (John et al. 2008) suggested how deployment is to be carried out for nodes used in the irrigation. (Diaz et al. 2011) addressed a methodology to monitor the agricultural production process based on soil moisture, humidity and luminosity for vineyard. (Liqiang et al. 2011) proposed a crop monitoring model measuring temperature and noting the reading for 24 hours in a day.

Sensors record data on a periodic basis and send them in a multi-hop fashion to a centralized processing unit. The centralized processing unit can be linked to external meteorological servers to get the global data useful for its computations. Data predicted can be used for forecasting, forewarning and ultimately to increase
productivity. Rain fed agriculture is totally based on rainfall pattern and related weather parameters. Monitoring the weather parameters regularly can assist in planning the cropping pattern accordingly. Reliable local metrological data and knowledge of soil moisture and ground water conditions can improve significantly agricultural water management. Sensor Network monitors the air temperature, humidity and ambient light intensity in a crop field and from remote places. It is observed that the placement of nodes, the measurement rate and the choice of mote depends on specific application.

2.4 Energy Computation Model

Energy consumed by the sensor is based on the distance between the sensors and also on the energy required to transmit and receive data. (Torres 2006) discussed the various energy computation models.

The classical energy computation model is proposed by (Heinzelman et al. 2000) where the transmitter and receiver electronics energy consumption are fixed to 50 nJ/bit. Transmitter amplifier consumes an additional power of 100 pJ/bit/m², while the idle and sleep power consumes 40 nJ/bit and 0 J respectively.

The μAMPS Specific Model uses transmission rate of 1 Mbps, so that one can calculate the energy required for transmitting one bit. The energy used in transmitting or receiving one bit is found by using the power value.

\[
\text{Time to send or receive one bit} = \text{bit/rate}=1 / 1 \text{ Mbps} = 1 \mu\text{sec} \tag{2.1}
\]

\[
\text{Energy} = \text{Power} \times \text{Time} \text{ where Power is in Watts and Time is in seconds.} \tag{2.2}
\]

The Power corresponds to the power consumption values from the data sheet for the active state of transmit and receive.

(Polastre et al. 2004) proposed the Mica2 Specific Model for measuring the energy consumption. The total energy E consumed by a node is given by,

\[
E = E_{rx} + E_{tx} + E_{listen} + E_d + E_{sleep} \tag{2.3}
\]
where $E_{rx}$ is the receive energy, $E_{tx}$ the transmit energy, $E_{listen}$ is the energy utilised when listening for messages on the radio channel, $E_d$ is the energy for sampling data energy and $E_{sleep}$ the energy consumed during the sleep period. 

$$E_d = t_d I_{data} V$$

where $t_d = t_{data} \times r$ where $r$ is the sampling frequency in per seconds. $t_d$ and $I_{data}$ represent the time to sample the data and current consumption for sampling. The lifetime $T_l$ of the node is dependent on the total energy consumed $E$, the battery capacity $C_{batt}$.

$$T_l = \frac{(C_{batt} \times V)}{(E \times 60 \times 60)} \text{ hours} \quad (2.4)$$

The procedure as proposed in the $\mu$AMPS Specific Model for energy computation and total energy using the Mica Specific Model are used in the computation of energy in the current work.

### 2.5 Connectivity Issues

The sensor network operates successfully if the active nodes maintain both sensing coverage and network connectivity. (Ghosh et al. 2008) suggested that using three sensors arranged along the boundary of the sensing range, provides minimal overlap with maximal coverage with minimum number of sensors. (Zhu et al. 2012) addressed that in WSN, connectivity is implied between sensors if the communicating range is twice the sensing range. The paper also summarizes the coverage and connectivity issues from three aspects: coverage deployment strategy, sleep scheduling mechanism and adjustable coverage radius. The optimal deployment pattern that achieves both coverage and $k$-connectivity (where $k$-connectivity implies $k$ nodes are in the communication range of each other) is obtained. (Zhang et al. 2005) proves that the condition of $rt \geq 2rs$ is both necessary and sufficient to ensure that complete coverage and connectivity is achieved in an arbitrary network of a convex region (where $rt$ is the transmission range and $rs$ the sensing range). It also addresses connectivity issues by keeping a minimum number of sensor nodes to operate in the active mode in wireless sensor network. (Li et al. 2009) determines the minimum density and optimal locations of relay and ordinary
sensors nodes to ensure connectivity, subject to various degrees of uncertainty in the locations of the nodes. (Hornsberg et al. 2006) proposed the grid structure for deployment with changing grid size with decreasing residual energy. (Younis et al. 2008) projected various strategies for node placement. (Wang et al. 2010; Penga et al. 2013; Al-karaki et al. 2009) addresses energy computation and data aggregation for the grid deployment. (Wang et al. 2011) proposed sub-dividing the hexagonal region into smaller regions to achieve longer network life. (Kaur and Baek 2009) suggested the usage of multiple sized fixed grid structure for the placement of nodes to reduce the consumption of energy. The placement of sensors nodes in the area of interest plays a vital role and directly affects the amount of energy consumed. It also decides on the coverage of the region and also if the nodes are connected to transmit and receive in the case of node failure.

2.6 Routing

One of the main design goals of WSN is to carry out data communication while trying to prolong the lifetime of the network and prevent connectivity degradation by employing aggressive energy management techniques. The design of routing protocols in WSN is influenced by many challenging factors. These factors must be overcome before efficient communication can be achieved in WSN. In the following, some of the routing challenges and design issues that affect the routing process in WSN are discussed. If the distribution of nodes is not uniform, optimal clustering becomes necessary to allow connectivity and enable energy-efficient network operation. In a multihop 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.

A number of survey papers on routing and aggregation in WSN are available. One of the earlier papers on survey of WSN is by (Akyildiz et al. 2002) classifying
the protocols based on routing. (Yick et al. 2008) addressed survey on WSN based on different proposed designs, algorithms, protocols, and services. In structure free mechanism random deployment of node is carried out. The structured based approach is classified into tree based, clustered approach and hybrid a combination of both of them. (Luo et al. 2007) classified routing schemes as correlated data gathering, routing driven algorithms with opportunistic aggregation, coding driven algorithm with partial aggregation, fusion driven algorithm with information reduction.

Based on the survey carried out, (Sohraby et al. 2007; Akkaya and Younis 2005) the broad classification of the routing protocols is formulated into network structure based or protocol structure based. Network structure is composed of flat, hierarchical and location based. In flat-based routing, all nodes are typically assigned equal roles or functionality. In hierarchical-based routing, nodes will play different roles in the network. In location-based routing, sensor nodes positions are exploited to route data in the network. Furthermore, these protocols can be classified into multipath-based, query-based, and negotiation-based, QoS-based, or coherent-based routing techniques depending on the protocol operation.

In addition to the above, routing protocols can be classified into three categories: proactive, reactive, and hybrid, depending on how the source finds a route to the destination. In proactive protocols, all routes are computed before they are really needed. Each network node stores information on its routes to all possible destinations at all times. When a routing request is needed, it can be initiated with minimal delay. Disadvantage is the overhead in the form of control messages that are required to establish and maintain all routes, even those that are infrequently used or subject to frequent failure.

In reactive protocols, routes are computed on demand. It requires less control information, when the network topology is rapidly changing or when there is data burst or is directed to only a small subset of the network nodes that are unknown in advance. Disadvantage is that, route requests can experience a large delay before
being propagated from the source node, as a route to the destination has to be determined first.

Hybrid protocols use a combination of these two ideas. When sensor nodes are static, it is preferable to have table-driven routing protocols rather than reactive protocols.

A significant amount of energy is used in route discovery and setup of reactive protocols. Another class of routing protocols is called cooperative. In cooperative routing, nodes send data to a central node where data can be aggregated and may be subject to further processing, hence reducing route cost in terms of energy use.

One of the earliest protocol Threshold sensitive Energy Efficient sensor Network (TEEN) is proposed by (Manjeshwar and Agrawal 2000) where the data is transmitted only if the changes in the data are above threshold. Power-Efficient GAthering in Sensor Information Systems (PEGASIS) is a chain based routing protocol suggested by (Lindsey et al. 2002). (Sharaf et al. 2005) addressed that group aware routing helps in reducing the number of data packets that can be sent to the base station. Quantile-digest by (Srivastava et al. 2004) provides information about the distribution of data, but not information concerning where those values occurred. (Mohanoor et al. 2009) projected a protocol based on the residual energy and minimum transmission cost (minimum hop count) along a path. (Jia et al. 2006) suggested a protocol based on minimum spanning tree. (Bari et al. 2009) addressed a genetic algorithm based routing protocol where in, a part of the route is swapped and new routes are established. (Li et al. 2008) established the Ant Colony Optimization process of finding the path from the source to destination. Several routes are discovered from the source to the destination and optimal route is found on the way back when path is established from the destination to the source. (Liu and Wang 2012) designed a protocol that not only led to balancing the energy expenditure among sensors, but also extending the network lifetime by equal usage of multiple optimal intermediate routers. (Misra and Thomasinous 2010) proposed
that the route to be taken should be such that it takes less time to reach the base station along with aggregation taking place at the first level of the tree.

Energy is consumed with each hop towards the destination for receiving and transmitting the packet. Reducing the number of hops reduces the energy consumption assisting in enhancing the life of the network. Various protocol designed give emphasis to minimum hop count, taking into account those nodes with higher node energy along the path. Data is transmitted from the location of the event occurrence to the destination in the case of event driven routing protocol. In environment monitoring applications all the sensors send data to the base station. The protocols used are mostly proactive in nature, where path from source to destination is established beforehand.

2.7 Clustering

A cluster is a subset of nodes from the network such that the data information is mutually reachable among each other within a region with single hop. Each cluster has a cluster- head and all the members are within direct radio range of the cluster-head. (Abbasi et al. 2007) proposed a survey on clustering techniques paying significant attention to clustering strategies and algorithms. Low-Energy Adaptive Clustering Hierarchy (LEACH) protocol by (Heinzelman et al. 2000) adopts a two-tier clustering structure, where information processing is performed at each cluster and routing is divided into two stages: routing from sensor nodes to cluster heads, and from cluster heads to the base station. Comparison of various protocols based on LEACH is carried out (Tyagi and Kumar 2013). (Song et al. 2007; Chao et al. 2010) proposed clustering along with consideration for energy savings. (Zhu et al. 2008) proposed the partitioning of network into rings along with creating sectors within each ring. This sector is changed dynamically to form cluster with different set of nodes and send data using minimum hop. (Liu et al. 2012) suggested that unequal clustering approach along with minimum hop inter cluster communication achieves minimum total energy consumption in the network. Energy aware
Hierarchical Agglomerative Clustering where in, repeated merging of small clusters are carried out until all clusters scale to the satisfied threshold and is addressed by (Du et al. 2013). (Latif et al. 2013) proposed dynamic cluster head selection for static clustering. (Yuea et al. 2012) divided network into unequal sized grid shaped clusters. The grid whose cluster head consumes more energy takes part in cluster head rotation and share energy load, balancing the energy dissipation. (Tang et al. 2012) proposed clustering such that some nodes send data in the form of chain to the cluster head which again aggregates at the sink. (Gao et al. 2010) suggested that the optimal cluster size is affected by the long-haul transmission distance and the nodal density of the network. (Jiang et al. 2011) proposed a method of clustering and sending of data based on the prediction of the value obtained in the sensor network. Clustering a set of nodes which have related information can assist in combining the data and reducing the number of times data transmission have to be carried out to the base station.

2.8 Energy Conservation Techniques

Various ways of energy conservation techniques are addressed by (Anastasi et al. 2009). In the current work the in-network processing along with data compression is carried out to conserve energy. Figure 2.2 shows a part of the various energy conservation schemes. Data-driven approaches can be divided according to the problem they address. Data reduction schemes address the case of unwanted samples, while energy-efficient data acquisition schemes are mainly aimed at reducing the energy spent by the sensing subsystem. However, some of them can reduce the energy spent for communication as well. All these techniques aim at reducing the amount of data to be delivered to the sink node. However the principles behind them are rather different.

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

![Energy Conservation Schemes Diagram]

*Fig. 2.2: A Sub Set of the Energy Saving Approaches in Sensor Network (Courtesy: (Anastasi et al. 2009))*

*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 decoding it at the sink.

*(Zhu et al. 2010)* showed that the minimal remaining energy and total energy consumption, plays a key role in prolonging network lifetime of a WSN. *(Deng et al. 2007)* suggested that a large transmission range increases the expected progress of a data packet towards its final destination at the expense of higher energy consumption per transmission. On the other hand, a short transmission range consumes less energy per transmission, but requires a larger number of hops for a data packet to reach its destination. *(Boukerche and Samarah 2009)* addressed
that data transmission can be reduced by using the association rules as in data mining.

A node can effectively increase its transmission range by reducing the data rate without changing its transmission power, contrary to fixing the data rate and adapting the transmission range by dynamically adjusting the transmission power. Furthermore, the interference among multiple traffic flows may affect the transmission range optimization by introducing packet collisions and retransmissions.

### 2.9 Aggregation

Data aggregation is the combination of data from different sources according to a certain aggregation function (e.g., duplicate suppression, minima, maxima, and average). 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 beamforming to combine the incoming signals and reducing the noise in these signals.

If every $k$ pieces of information can be aggregated into a single piece of information, the routing can be aggregated to a single piece of information. This reduces the routing load by a factor of at most $k$. The value of $k$ is usually referred to as the *aggregation factor*, or *correlation factor* among data.

Early work on data centric routing as in directed diffusion (Intanagonwiwat et al. 2003) and Sensor Protocols for Information via Negotiation (SPIN) by (Rabiner et al. 1999) is shown to save energy through elimination of redundant data and data negotiation. In SPIN, topological changes are localized since each node need know only its single-hop neighbours. It provides more energy savings than flooding, and metadata negotiation almost halves the redundant data. (Zhou et al. 2006) aims at improving the Directed Diffusion algorithm by avoiding unnecessary broadcasting of the control signal using hierarchical structure. Hierarchical routing
is an efficient way to lower energy consumption within a cluster, performing data aggregation and fusion in order to decrease the number of transmitted messages to the base station.

(Luo et al. 2007; Rajagopalan et al. 2006) suggested different aggregation techniques focusing on optimizing important performance measures such as network lifetime, data latency, data accuracy and energy consumption. (Fasolo et al. 2007) summarised the basic characteristics of the data aggregation functions: *lossy, duplicate sensitive* and *correlated* aware. (Nakamura et al. 2007) brings the methods, models and classifications of information fusion. Using Tiny Aggregation (TaG) protocol (Madden et al. 2002) demonstrated the various aggregation functions like count, max. /min., sum, average, median, count distinct and histogram. (Intanagonwiwat et al. 2002) demonstrated the advantage of early aggregation on the energy conservation.

(Kamat et al. 2011) proposed the fine grained aggregation where wedges are created in the network and data is aggregated from the outer hop to the inner hop aggregating along the way to the base station. (Sutagundar et al. 2013) proposed the data aggregation along the spokes of the wheel. (Lu et al. 2012) proposed a ring based data gathering for WSN. Combining tree and multipath concept is implemented by (Manjhi A. et al. 2005). An energy-aware spatial correlation mechanism in which nodes which detect the same event are grouped in correlated regions and a representative node is selected at each correlation region dynamically. Dynamic and scalable tree which is aware of spatial correlation is introduced by (Leandro et al. 2011). Best response dynamics to local data is discussed (Zeydan et al. 2012). (Motegi et al. 2006) proposed how in-network aggregation can be carried out with multiple parents. (Fan et al. 2006) proposed a Tree on Directed Acyclic Graph where the clustering and aggregation takes place on demand. This avoids the long stretch problem in fixed structured approaches and eliminates the overhead of construction and maintenance of dynamic structures. (Villas et al. 2011) proposed a grid-based routing and aggregator selection scheme
so as to achieve low energy dissipation and low latency by finding the minimum number of aggregation points while routing data to the base station. (Ding et al. 2003) proposed protocol such that the leaf node is put to sleep when not in use to conserve energy.

2.10 Simulators

In order to evaluate the behavior and performance of protocols for wireless networks, simulations are a good compromise between cost and complexity, on the one hand, and accuracy of the results, on the other hand. Since there are many simulators for wireless networks, it is often difficult to decide which simulator to choose. Few among them are Java Simulator (J-Sim), Optimized Network Engineering Tools (OPNET), nesC with TinyOS Simulator (TOSSIM), Objective Modular Network Testbed with C++ (OMNET++), Global Mobile Information System Simulator (GlomoSim), Sensor Environment and Network Simulator (SENS). There are large numbers of directly supportive sensor network simulators. SENS modules are programmed in C++ and TOSSIM whose TinyOS codes are used to support the high level sensor network simulation. But TOSSIM is not sufficient for the low level protocol such as medium access control (MAC). OPNeT++ is programmed in C++ while its compound modules are programmed in a high-level language (NED). Although J-Sim provides supporting target, sensor and sink nodes, sensor channels and wireless communication channels, its use of Java as the simulating language is inevitably sacrificing the efficiency of the simulation. As the packet formats, energy models and MAC protocols are not representative of those used in wireless sensor networks, GloMoSim may not be a good option for the sensor network simulator. OMNET++ is another tool used for WSN. (Korkalainen et al. 2009) proposed a case study of various simulators. Networked system using C (nesC) programming model incorporates event-driven execution, a flexible concurrency model, and component-oriented application design. TinyOS is an event-driven operating system intended for sensor networks with limited resources.
The TinyOS system, libraries, and applications are written in nesC, a new language for programming embedded systems. nesC is a version of C that is designed to handle the structuring and execution of TinyOS.

The network simulator (ns-2) which supports both wired and wireless networks is also implemented to be used with the wireless sensor network. This ns-2 based simulator has been used to test various routing protocols. Naval Research Laboratory (NRL) Sensor Sim as proposed by (Downard 2004) is an extension to ns-2 to support sensor network. Mannasim is another framework which supports WSN. (Ros and Ruiz 2004; Talipov et al. 2010) explained the implementation of new routing protocol. Based on certain features like the popularity of the protocol, the availability of patches and freely downloadable versions of ns-2 and its extensions, ns-2 is used in the proposed work.

2.11 Summary

In this Chapter an insight into the domain knowledge, of sensor network in agriculture is studied. Heavy use of groundwater leading to steady declines in water tables is a problem increasingly witnessed in many parts of the world. In order to cover wide areas and to monitor efficiently water resources, the use of several sensors working in a network seems particularly appropriate. Common-Sense Net consists in a wireless network of ground-sensors that record periodically the state of the soil. The study of CSN has laid the foundation and motivation for the development of the proposed system.

The chapter also focuses on various routing, clustering, aggregation and energy conservation and computation techniques. Energy conservation is the main goal of the work. To achieve this, routing techniques with minimal hop count, proper clustering and appropriate aggregation are essential so as to minimize the number of transmission and reception of data packets.
The literature survey suggests that opportunistic aggregation takes place along the path. The algorithms specified in the literature do not work on the actual data. In the current work as ns-2 does not support data handling, the simulator is extended to support data at the application and the transport layer. As the literature surveyed does not specify handling of data, no paper discusses about the number of nodes holding same data to carry out aggregation. This counter is essential to carry out average of the values at a particular area.

Numerous challenges make the study of real deployed sensor networks very difficult and financially infeasible. At the current stage of the technology, a practical way to study WSN is through simulations that can provide a meaningful perspective of the behavior and performance of various algorithms.