

Wireless Sensor Networks¹

F. L. LEWIS

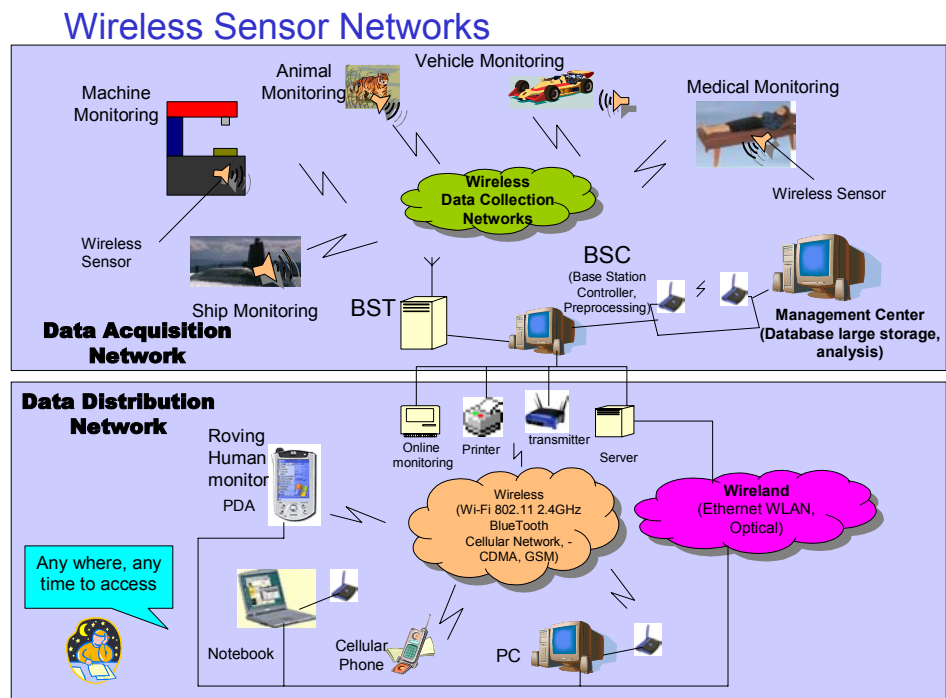
Associate Director for Research
Head, Advanced Controls, Sensors, and MEMS Group
Automation and Robotics Research Institute
The University of Texas at Arlington
7300 Jack Newell Blvd. S
Ft. Worth, Texas 76118-7115
email lewis@uta.edu, <http://arri.uta.edu/acs>

2.1. INTRODUCTION

Smart environments represent the next evolutionary development step in building, utilities, industrial, home, shipboard, and transportation systems automation. Like any sentient organism, the smart environment relies first and foremost on sensory data from the real world. Sensory data comes from multiple sensors of different modalities in distributed locations. The smart environment needs information about its surroundings as well as about its internal workings; this is captured in biological systems by the distinction between *exteroceptors* and *proprioceptors*.

The challenges in the hierarchy of: detecting the relevant quantities, monitoring and collecting the data, assessing and evaluating the information, formulating meaningful user displays, and performing decision-making and alarm functions are enormous. The information needed by smart environments is provided by Distributed Wireless Sensor Networks, which are responsible for sensing as well as for the first stages of the processing hierarchy. The importance of sensor networks is highlighted by the number of recent funding initiatives, including the DARPA SENSIT program, military programs, and NSF Program Announcements.

The figure shows the complexity of wireless sensor networks, which generally consist of a data acquisition network and a data distribution network, monitored and controlled by a management center. The plethora of available technologies makes even the selection of



¹ This research was supported by ARO Research Grant DAAD 19-02-1-0366

components difficult, let alone the design of a consistent, reliable, robust overall system.

The study of wireless sensor networks is challenging in that it requires an enormous breadth of knowledge from an enormous variety of disciplines. In this chapter we outline communication networks, wireless sensor networks and smart sensors, physical transduction principles, commercially available wireless sensor systems, self-organization, signal processing and decision-making, and finally some concepts for home automation.

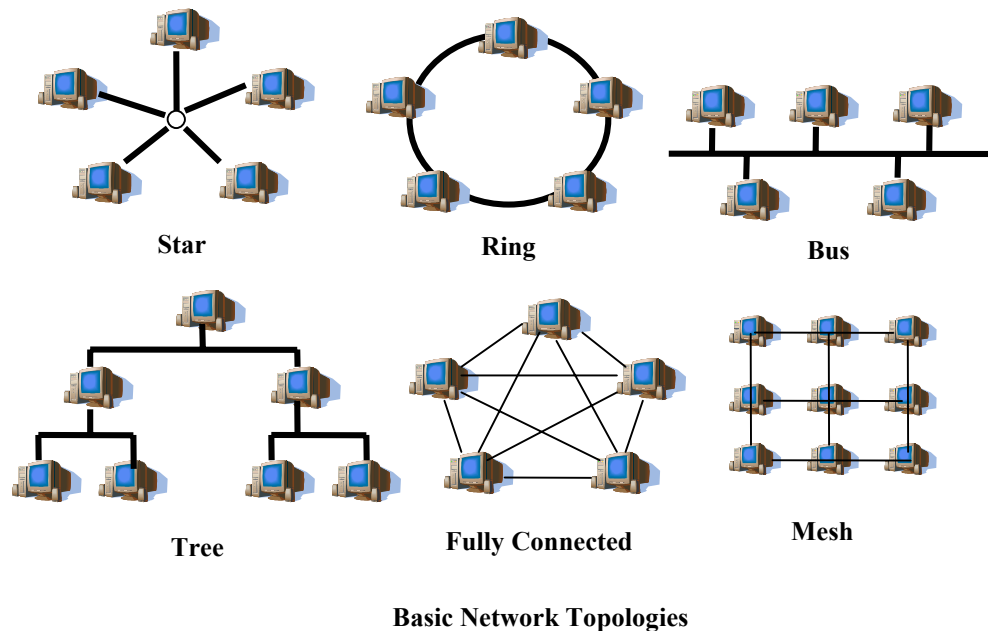
2.2. COMMUNICATION NETWORKS

The study of communication networks can encompass several years at the college or university level. To understand and be able to implement sensor networks, however, several basic primary concepts are sufficient.

2.2.1. Network Topology

The basic issue in communication networks is the transmission of messages to achieve a prescribed message throughput (Quantity of Service) and Quality of Service (QoS). QoS can be specified in terms of message delay, message due dates, bit error rates, packet loss, economic cost of transmission, transmission power, etc. Depending on QoS, the installation environment, economic considerations, and the application, one of several basic network topologies may be used.

A communication network is composed of nodes, each of which has computing power and can transmit and receive messages over communication links, wireless or cabled. The basic network topologies are shown in the figure and include fully connected, mesh, star, ring, tree, bus. A single network may consist of several interconnected subnets of different topologies. Networks are further classified as Local Area Networks (LAN), e.g. inside one building, or Wide Area Networks (WAN), e.g. between buildings.



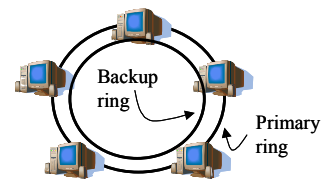
Fully connected networks suffer from problems of NP-complexity [Garey 1979]; as additional nodes are added, the number of links increases exponentially. Therefore, for large networks, the routing problem is computationally intractable even with the availability of large amounts of computing power.

Mesh networks are regularly distributed networks that generally allow transmission only to a node's nearest neighbors. The nodes in these networks are generally identical, so that mesh nets are also referred to as peer-to-peer (see below) nets. Mesh nets can be good models for large-scale networks of wireless sensors that are distributed over a geographic region, e.g. personnel or vehicle security surveillance systems. Note that the regular structure reflects the communications topology; the actual geographic distribution of the nodes need not be a regular mesh. Since there are generally multiple routing paths between nodes, these nets are robust to failure of individual nodes or links. An advantage of mesh nets is that, although all nodes may be identical and have the same computing and transmission capabilities, certain nodes can be designated as 'group leaders' that take on additional functions. If a group leader is disabled, another node can then take over these duties.

All nodes of the **star topology** are connected to a single hub node. The hub requires greater message handling, routing, and decision-making capabilities than the other nodes. If a communication link is cut, it only affects one node. However, if the hub is incapacitated the network is destroyed. In the **ring topology** all nodes perform the same function and there is no leader node. Messages generally travel around the ring in a single direction.

However, if the ring is cut, all communication is lost. The **self-healing ring network** (SHR) shown has two rings and is more fault tolerant.

In the **bus topology**, messages are broadcast on the bus to all nodes. Each node checks the destination address in the message header, and processes the messages addressed to it. The bus topology is passive in that each node simply listens for messages and is not responsible for retransmitting any messages.



Self-Healing Ring

2.2.2. Communication Protocols and Routing

The topics of communication protocols and routing are complex and require much study. Some basics useful for understanding sensor nets are presented here.

Headers. Each message generally has a *header* identifying its source node, destination node, length of the data field, and other information. This is used by the nodes in proper routing of the message. In encoded messages, parity bits may be included. In *packet routing networks*, each message is broken into *packets* of fixed length. The packets are transmitted separately through the network and then reassembled at the destination. The fixed packet length makes for easier routing and satisfaction of QoS. Generally, voice communications use circuit switching, while data transmissions use packet routing.

Preamble 8 bytes	Destination Address 6 bytes	Source Address 6 bytes	Length of Data field 2 bytes	Protocol header, Data, padding 0-1500 bytes
---------------------	-----------------------------------	------------------------------	------------------------------------	---

Ethernet Message Header

In addition to the information content messages, in some protocols (e.g. FDDI- see below) the nodes transmit *special frames* to report and identify fault conditions. This can allow network reconfiguration for fault recovery. Other special frames might include route discovery packets or *ferrets* that flow through the network, e.g. to identify shortest paths, failed links, or transmission cost information. In some schemes, the ferret returns to the source and reports the best path for message transmission.

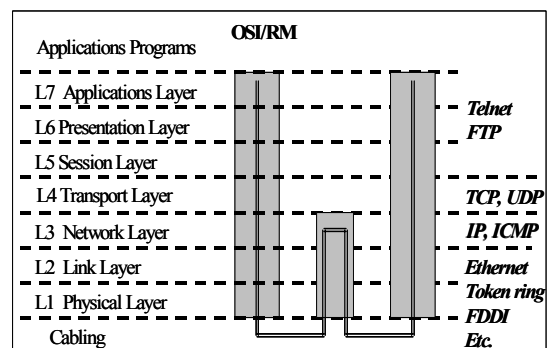
When a node desires to transmit a message, *handshaking protocols* with the destination node are used to improve reliability. The source and destination might transmit alternately as follows: request to send, ready to receive, send message, message received. Handshaking is used to guarantee QoS and to retransmit messages that were not properly received.

Switching. Most computer networks use a *store-and-forward* switching technique to control the flow of information [Duato 1996]. Then, each time a packet reaches a node, it is completely buffered in local memory, and transmitted as a whole. More sophisticated switching techniques include *wormhole*, which splits the message into smaller units known as flow control units or flits. The header flit determines the route. As the header is routed, the remaining flits follow it in pipeline fashion. This technique currently achieves the lowest message latency. Another popular switching scheme is *virtual-cut-through*. Here, when the header arrives at a node, it is routed without waiting for the rest of the packet. Packets are buffered either in software buffers in memory or in hardware buffers, and various sorts of buffers are used including edge buffers, central buffers, etc.

Multiple Access Protocols. When multiple nodes desire to transmit, protocols are needed to avoid collisions and lost data. In the ALOHA scheme, first used in the 1970's at the University of Hawaii, a node simply transmits a message when it desires. If it receives an acknowledgement, all is well. If not, the node waits a random time and retransmits the message.

In *Frequency Division Multiple Access* (FDMA), different nodes have different carrier frequencies. Since frequency resources are divided, this decreases the bandwidth available for each node. FDMA also requires additional hardware and intelligence at each node. In *Code Division Multiple Access* (CDMA), a unique code is used by each node to encode its messages. This increases the complexity of the transmitter and the receiver. In *Time Division Multiple Access* (TDMA), the RF link is divided on a time axis, with each node being given a predetermined time slot it can use for communication. This decreases the sweep rate, but a major advantage is that TDMA can be implemented in software. All nodes require accurate, synchronized clocks for TDMA.

Open Systems Interconnection Reference Model (OSI/RM). The International Standards Organization (ISO) OSI/RM architecture specifies the relation between messages transmitted in a communication network and applications



Open Systems Interconnection Reference Model

programs run by the users. The development of this open standard has encouraged the adoption by different developers of standardized compatible systems interfaces. The figure shows the seven layers of OSI/RM. Each layer is self-contained, so that it can be modified without unduly affecting other layers. The Transport Layer provides error detection and correction. Routing and flow control are performed in the Network Layer. The Physical Layer represents the actual hardware communication link interconnections. The Applications Layer represents programs run by users.

Routing. Since a distributed network has multiple nodes and services many messages, and each node is a shared resource, many decisions must be made. There may be multiple paths from the source to the destination. Therefore, message routing is an important topic. The main performance measures affected by the routing scheme are throughput (quantity of service) and average packet delay (quality of service). Routing schemes should also avoid both deadlock and livelock (see below).

Routing methods can be fixed (i.e. pre-planned), adaptive, centralized, distributed, broadcast, etc. Perhaps the simplest routing scheme is the *token ring* [Smythe 1999]. Here, a simple topology and a straightforward fixed protocol result in very good reliability and precomputable QoS. A token passes continuously around a ring topology. When a node desires to transmit, it captures the token and attaches the message. As the token passes, the destination reads the header, and captures the message. In some schemes, it attaches a ‘message received’ signal to the token, which is then received by the original source node. Then, the token is released and can accept further messages. The token ring is a completely decentralized scheme that effectively uses TDMA. Though this scheme is very reliable, one can see that it results in a waste of network capacity. The token must pass once around the ring for each message. Therefore, there are various modifications of this scheme, including using several tokens, etc.

Fixed routing schemes often use Routing Tables that dictate the next node to be routed to, given the current message location and the destination node. Routing tables can be very large for large networks, and cannot take into account real-time effects such as failed links, nodes with backed up queues, or congested links.

Adaptive routing schemes depend on the current network status and can take into account various performance measures, including cost of transmission over a given link, congestion of a given link, reliability of a path, and time of transmission. They can also account for link or node failures.

Routing algorithms can be based on various network analysis and graph theoretic concepts in Computer Science (e.g. A-star tree search), or in Operations Research [Bronson 1997] including shortest-route, maximal flow, and minimum-span problems. Routing is closely associated with dynamic programming and the optimal control problem in feedback control theory [Lewis and Syrmos 1995]. Shortest Path routing schemes find the shortest path from a given node to the destination node. If the cost, instead of the link length, is associated with each link, these algorithms can also compute minimum cost routes. These algorithms can be centralized (find the shortest path from a given node to all other nodes) or decentralized (find the shortest path from all nodes to a given node). There are certain well-defined algorithms for shortest path routing, including the efficient Dijkstra algorithm [Kumar 2001], which has polynomial complexity. The Bellman-Ford algorithm finds the path with the least number of hops [Kumar 2001]. Routing schemes based on competitive game theoretic notions have also been developed [Altman et al. 2002].

Deadlock and Livelock. Large-scale communication networks contain cycles (circular paths) of nodes. Moreover, each node is a shared resource that can handle multiple messages flowing along different paths. Therefore, communication nets are susceptible to *deadlock*, wherein all nodes in a specific cycle have full buffers and are waiting for each other. Then, no node can transmit because no node can get free buffer space, so all transmission in that cycle comes to a halt. *Livelock*, on the other hand, is the condition wherein a message is continually transmitted around the network and never reaches its destination. Livelock is a deficiency of some routing schemes that route the message to alternate links when the desired links are congested, without taking into account that the message should be routed closer to its final destination. Many routing schemes are available for routing with deadlock and livelock avoidance [e.g. Duato 1996].

Flow Control. In queuing networks, each node has an associated queue or buffer that can stack messages. In such networks, flow control and resource assignment are important. The objectives of flow control are to protect the network from problems related to overload and speed mismatches, and to maintain QoS, efficiency, fairness, and freedom from deadlock. If a given node *A* has high priority, its messages might be preferentially routed in every case, so that competing nodes are choked off as the traffic of *A* increases. Fair routing schemes avoid this. There are several techniques for flow control: In *buffer management*, certain portions of the buffer space are assigned for certain purposes. In *choke packet schemes*, any node sensing congestion sends choke packets to other nodes telling them to reduce their transmissions. *Isarithmic schemes* have a fixed number of ‘permits’ for the network. A message can be sent only if a permit is available. In *window* or *kanban* schemes, the receiver grants ‘credits’ to the sender only if it has free buffer space. Upon receiving a credit, the sender can transmit a message. In *Transmission*

Control Protocol (TCP) schemes (Tahoe and Reno) a source linearly increases its transmission rate as long as all its sent messages are acknowledged for. When it detects a lost packet, it exponentially decreases its transmission rate. Since lost packets depend on congestion, TCP automatically decreases transmissions when congestion is detected.

2.2.3. Power Management

With the advent of ad hoc networks of geographically distributed sensors in remote site environments (e.g. sensors dropped from aircraft for personnel/vehicle surveillance), there is a focus on increasing the lifetimes of sensor nodes through power generation, power conservation, and power management. Current research is in designing small MEMS (microelectromechanical systems) RF components for transceivers, including capacitors, inductors, etc. The limiting factor now is in fabricating micro-sized inductors. Another thrust is in designing MEMS power generators using technologies including solar, vibration (electromagnetic and electrostatic), thermal, etc.

RF-ID (RF identification) devices are transponder microcircuits having an L-C tank circuit that stores power from received interrogation signals, and then uses that power to transmit a response. Passive tags have no onboard power source and limited onboard data storage, while active tags have a battery and up to 1Mb of data storage. RF-ID operates in a low frequency range of 100kHz-1.5MHz or a high frequency range of 900 MHz-2.4GHz, which has an operating range up to 30m. RF-ID tags are very inexpensive, and are used in manufacturing and sales inventory control, container shipping control, etc. RF-ID tags are installed on water meters in some cities, allowing a metering vehicle to simply drive by and remotely read the current readings. They are also be used in automobiles for automatic toll collection.

Meanwhile, software power management techniques can greatly decrease the power consumed by RF sensor nodes. TDMA is especially useful for power conservation, since a node can power down or ‘sleep’ between its assigned time slots, waking up in time to receive and transmit messages.

The required transmission power increases as the square of the distance between source and destination. Therefore, multiple short message transmission hops require less power than one long hop. In fact, if the distance between source and destination is R , the power required for single-hop transmission is proportional to R^2 . If nodes between source and destination are taken advantage of to transmit n short hops instead, the power required by each node is proportional to R^2/n^2 . This is a strong argument in favor of distributed networks with multiple nodes, i.e. nets of the mesh variety.

A current topic of research is *active power control*, whereby each node cooperates with all other nodes in selecting its individual transmission power level [Kumar 2001]. This is a decentralized feedback control problem. Congestion is increased if any node uses too much power, but each node must select a large enough transmission range that the network remains connected. For n nodes randomly distributed in a disk, the network is asymptotically connected with probability one if the transmission range r of all nodes is selected using

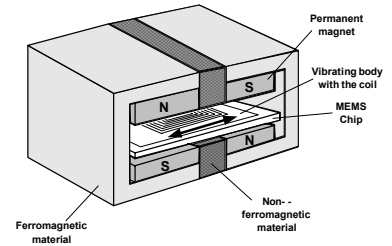
$$r \geq \sqrt{\frac{\log n + \gamma(n)}{\pi n}}$$

where $\gamma(n)$ is a function that goes to infinity as n becomes large.

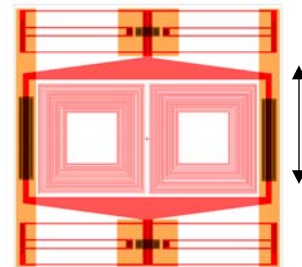
2.2.4. Network Structure and Hierarchical Networks

Routing tables for distributed networks increase exponentially as nodes are added. An $n \times m$ mesh network has nm links, and there are multiple paths from each source to each destination. *Hierarchical network structures* simplify routing, and also are amenable to distributed signal processing and decision-making, since some processing can be done at each hierarchical layer.

It has been shown [Lewis and Abdallah 1993] that a fully connected network has *NP-hard* complexity, while imposing routing protocols by restricting the allowed paths to obtain a *reentrant flow* topology results in polynomial complexity. Such streamlined protocols are natural for hierarchical networks.

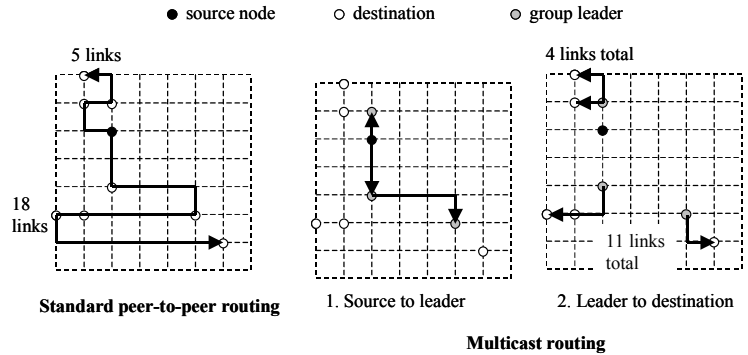


MEMS power generator using vibration and electromagnetic method



MEMS fabrication layout of power generator dual vibrating coil showing folded beam suspension.

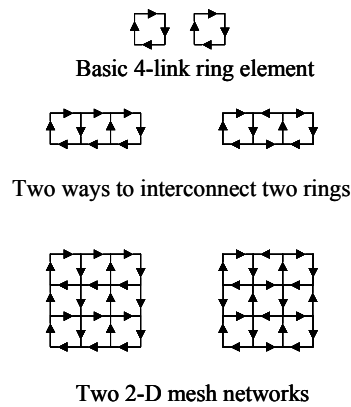
Multicast Systems in mesh networks use a hierarchical leader-based scheme for message transmission [Chen et al. 2000]. Each group of nodes has a designated leader that is responsible for receiving messages from and transmitting to nodes outside the group. Part (a) of the figure shows messages routed in a mesh net using standard peer-to-peer protocols. The link lengths of the transmission paths are shown. Parts (b) and (c) show the same two messages being routed using a multicast protocol. Note that the total transmission paths are significantly shorter. Multicast has been implemented using tree-based and path-based schemes.



Taken from Chen et al. (2000)

Multicast routing improves efficiency and reduces message path length

Hierarchical Networks. Much work has been done on formal hierarchical structures for distributed networks. Cao [1999] studies how to determine optimal configurations for hierarchical routing. Shi [1995] analyzes hierarchical self-healing rings. Shah-Heydari [2001] shows the importance of a consistent numbering scheme in hierarchical systems, which allows for a simplified tree-based routing scheme.



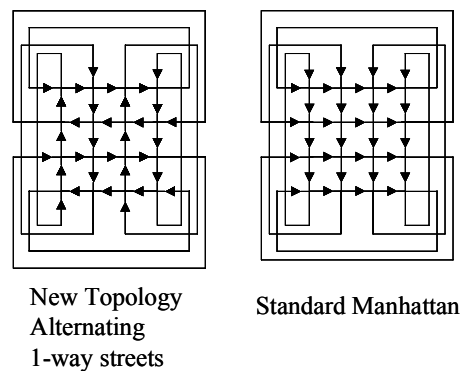
Constructing two mesh networks

The figure shows a basic 4-link ring element consisting of four nodes and four links. It shows two ways of connecting these two rings, which results in two mesh networks of different structures. The first network consists of alternating one-way streets, while the second consists of alternating-direction vortices. It is interesting to analyze these two structures from the point of view of the notions of flow field divergence and curl.

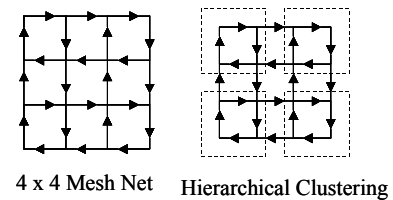
In any network, the phenomenon of *edge binding* means that much of the routing power of peripheral stations is wasted because peripheral links are unused. Thus, messages tend to reflect off the boundary into the interior or to move parallel to the periphery [J.W. Smith, Rand Corp. 1964]. To avoid this, the Manhattan geometry connects the nodes at one edge of the network to nodes at the opposite edge. The figure shows the standard Manhattan geometry as well as a Manhattan net built from the alternating one-way street mesh just constructed.

As nodes are added, the number of links increases exponentially. This makes for NP-complexity problems in routing and failure recovery. To simplify network structure, we can use hierarchical clustering techniques. The hierarchical structure must be *consistent*, that is, it must have the same structure at each level. The figure shows a 4x4 mesh net and also a clustering into four groups. Note that the clustered structure has a dual ring SHR topology. To reduce the routing complexity, we can disable one of the rings and obtain a ring structure.

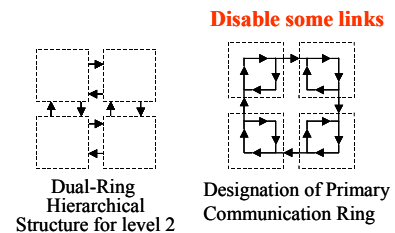
The next figure shows an 8x8 mesh net. Shown first are all the links, and then the hierarchical clustering with some links disabled to reduce complexity. We have chosen to keep the outer ring at each level. Note that the clockwise ring structure is the same at each level, resulting in a regular hierarchy.



Interconnecting the edge links

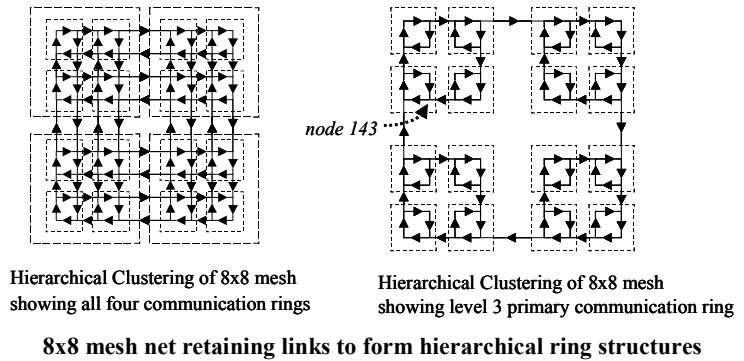


Clustering the nodes



Reducing complexity

Routing is very easy in this hierarchical network [Swamy 2003]. First, one selects a *consistent numbering scheme*. For example number the groups as 1,2,3,4 beginning in the top left and going clockwise. This is done at each level. Then, referring to the 8x8 mesh net in the figure, node 143, shown in the figure, is in the top left 4x4 group, within which it is in the fourth 2x2 group, within which it is the third node. Using this number scheme one may construct a simple routing scheme wherein the same basic routing algorithm is repeated at each level of the hierarchy. This is not unlike quadtree routing in mobile robot path planning. Failure recovery is also straightforward. If a link fails, one may simply switch in one of the disabled links to take over. Code for this is very easy to write.



Distributed Routing, Decision-Making, and DSP. It is natural in routing and failure recovery for these hierarchical networks to designate the entry node for each group as a *group leader*. This node must make additional decisions beyond those of the other nodes, including resource availability for deadlock avoidance, disabled link activation for failure recovery, and so on. This lays a very natural framework for distributed decision-making and digital signal processing (DSP), wherein a group leader processes the data from the group prior to transmitting it. The group leader for communications should be the *entry node* of each group, while the group leader for DSP should be the *exit node* for each group.

2.2.5. Historical Development and Standards

Much of this information is taken from [PC Tech Guide], which contains a thorough summary of communication network standards, topologies, and components. See also Jordan and Abdallah [2002].

Ethernet. The Ethernet was developed in the mid 1970's by Xerox, DEC, and Intel, and was standardized in 1979. The Institute of Electrical and Electronics Engineers (IEEE) released the official Ethernet standard IEEE 802.3 in 1983. The Fast Ethernet operates at ten times the speed of the regular Ethernet, and was officially adopted in 1995. It introduces new features such as full-duplex operation and auto-negotiation. Both these standards use IEEE 802.3 variable-length frames having between 64 and 1514-byte packets.

Token Ring. In 1984 IBM introduced the 4Mbit/s token ring network. The system was of high quality and robust, but its cost caused it to fall behind the Ethernet in popularity. IEEE standardized the token ring with the IEEE 802.5 specification. The Fiber Distributed Data Interface (FDDI) specifies a 100Mbit/s token-passing, dual-ring LAN that uses fiber optic cable. It was developed by the American National Standards Institute (ANSI) in the mid 1980s, and its speed far exceeded current capabilities of both Ethernet and IEEE 802.5.

Gigabit Ethernet. The Gigabit Ethernet Alliance was founded in 1996, and the Gigabit Ethernet standards were ratified in 1999, specifying a physical layer that uses a mixture of technologies from the original Ethernet and fiber optic cable technologies from FDDI.

Client-Server networks became popular in the late 1980's with the replacement of large mainframe computers by networks of personal computers. Application programs for distributed computing environments are essentially divided into two parts: the client or front end, and the server or back end. The user's PC is the client and more powerful server machines interface to the network.

Peer-to-Peer networking architectures have all machines with equivalent capabilities and responsibilities. There is no server, and computers connect to each other, usually using a bus topology, to share files, printers, Internet access, and other resources.

Peer-to-Peer Computing is a significant next evolutionary step over P2P networking. Here, computing tasks are split between multiple computers, with the result being assembled for further consumption. P2P computing has sparked a revolution for the Internet Age and has obtained considerable success in a very short time. The Napster MP3 music file sharing application went live in September 1999, and attracted more than 20 million users by mid 2000.

802.11 Wireless Local Area Network. IEEE ratified the IEEE 802.11 specification in 1997 as a standard for WLAN. Current versions of 802.11 (i.e. 802.11b) support transmission up to 11Mbit/s. WiFi, as it is known, is useful for fast and easy networking of PCs, printers, and other devices in a local environment, e.g. the home. Current PCs and laptops as purchased have the hardware to support WiFi. Purchasing and installing a WiFi router and receivers is within the budget and capability of home PC enthusiasts.

Bluetooth was initiated in 1998 and standardized by the IEEE as Wireless Personal Area Network (WPAN) specification IEEE 802.15. Bluetooth is a short range RF technology aimed at facilitating communication of electronic devices between each other and with the Internet, allowing for data synchronization that is transparent to the user. Supported devices include PCs, laptops, printers, joysticks, keyboards, mice, cell phones, PDAs, and consumer products. Mobile devices are also supported. Discovery protocols allow new devices to be hooked up easily to the network. Bluetooth uses the unlicensed 2.4 GHz band and can transmit data up to 1Mbit/s, can penetrate solid non-metal barriers, and has a nominal range of 10m that can be extended to 100m. A master station can service up to 7 simultaneous slave links. Forming a network of these networks, e.g. a piconet, can allow one master to service up to 200 slaves.

Currently, Bluetooth development kits can be purchased from a variety of suppliers, but the systems generally require a great deal of time, effort, and knowledge for programming and debugging. Forming piconets has not yet been streamlined and is unduly difficult.

Home RF was initiated in 1998 and has similar goals to Bluetooth for WPAN. Its goal is shared data/voice transmission. It interfaces with the Internet as well as the Public Switched Telephone Network. It uses the 2.4 GHz band and has a range of 50 m, suitable for home and yard. A maximum of 127 nodes can be accommodated in a single network. **IrDA** is a WPAN technology that has a short-range, narrow-transmission-angle beam suitable for aiming and selective reception of signals.

2.3. WIRELESS SENSOR NETWORKS

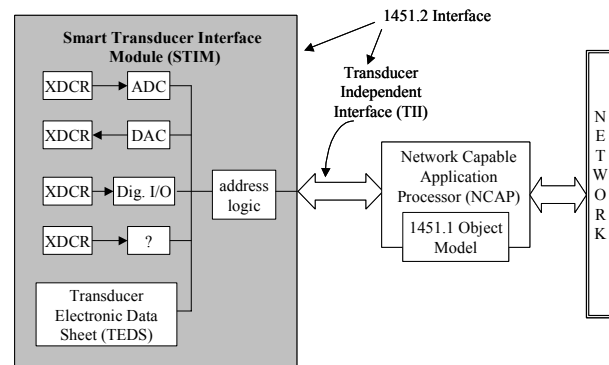
Sensor networks are the key to gathering the information needed by smart environments, whether in buildings, utilities, industrial, home, shipboard, transportation systems automation, or elsewhere. Recent terrorist and guerilla warfare countermeasures require distributed networks of sensors that can be deployed using, e.g. aircraft, and have self-organizing capabilities. In such applications, running wires or cabling is usually impractical. A sensor network is required that is fast and easy to install and maintain.

2.3.1. IEEE 1451 and Smart Sensors

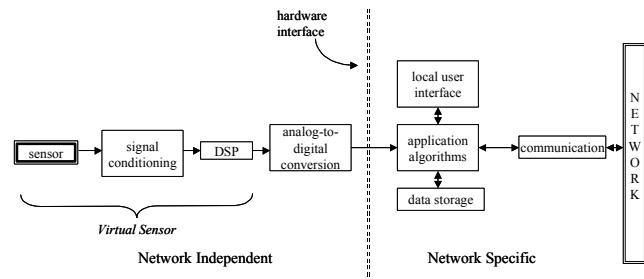
Wireless sensor networks satisfy these requirements. Desirable functions for sensor nodes include: ease of installation, self-identification, self-diagnosis, reliability, time awareness for coordination with other nodes, some software functions and DSP, and standard control protocols and network interfaces [IEEE 1451 Expo, 2001].

There are many sensor manufacturers and many networks on the market today. It is too costly for manufacturers to make special transducers for every network on the market. Different components made by different manufacturers should be compatible. Therefore, in 1993 the IEEE and the National Institute of Standards and Technology (NIST) began work on a standard for Smart Sensor Networks. IEEE 1451, the Standard for Smart Sensor Networks was the result. The objective of this standard is to make it easier for different manufacturers to develop smart sensors and to interface those devices to networks.

Smart Sensor, Virtual Sensor. The figure shows the basic architecture of IEEE 1451 [Conway and Hefferman 2003]. Major components include STIM, TEDS, TII, and NCAP as detailed in the figure. A major outcome of IEEE 1451 studies is the formalized concept of a *Smart Sensor*. A smart sensor is a sensor that provides extra functions beyond those necessary for generating a correct representation of the sensed quantity [Frank 2000]. Included might be signal conditioning, signal processing, and decision-making/alarm functions. A general model of a smart sensor is shown in the figure. Objectives for smart sensors include moving the intelligence closer to the point of measurement; making it cost effective to integrate and maintain distributed sensor systems; creating a



The IEEE 1451 Standard for Smart Sensor Networks

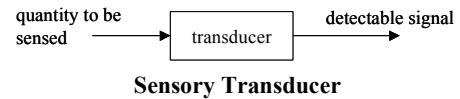


A general model of a smart sensor [IEEE 1451 Expo, Oct. 2001]

confluence of transducers, control, computation, and communications towards a common goal; and seamlessly interfacing numerous sensors of different types. The concept of a *Virtual Sensor* is also depicted. A virtual sensor is the physical sensor/transducer, plus the associated signal conditioning and digital signal processing (DSP) required to obtain reliable estimates of the required sensory information. The virtual sensor is a component of the smart sensor.

2.3.2. Transducers and Physical Transduction Principles

A *transducer* is a device that converts energy from one domain to another. In our application, it converts the quantity to be sensed into a useful signal that can be directly measured and processed. Since much signal conditioning (SC) and digital signal processing (DSP) is carried out by electronic circuits, the outputs of transducers that are useful for sensor networks are generally voltages or currents. Sensory transduction may be carried out using *physical principles*, some of which we review here. Microelectromechanical Systems (MEMS) sensors are by now very well developed and are available for most sensing applications in wireless networks. References for this section include Frank [2000], Kovacs [1998], Madou [1997], de Silva [1999].



Mechanical Sensors include those that rely on direct physical contact.

The Piezoresistive Effect converts an applied strain to a change in resistance that can be sensed using electronic circuits such as the Wheatstone Bridge (discussed later). Discovered by Lord Kelvin in 1856, the relationship is $\Delta R/R = S\varepsilon$, with R the resistance, ε the strain, and S the gauge factor which depends on quantities such as the resistivity and the Poisson's ratio of the material. There may be a quadratic term in ε for some materials. Metals and semiconductors exhibit piezoresistivity. The piezoresistive effect in silicon is enhanced by doping with boron (p-type silicon can have a gauge factor up to 200). With semiconductor strain gauges, temperature compensation is important.

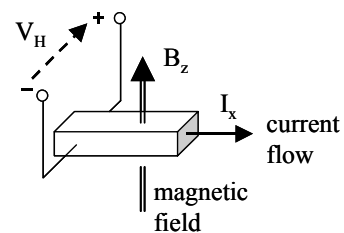
The Piezoelectric Effect, discovered by the Curies in 1880, converts an applied stress (force) to a charge separation or potential difference. Piezoelectric materials include barium titanate, PZT, and single-crystal quartz. The relation between the change in force F and the change in voltage V is given by $\Delta V = k \Delta F$, where k is proportional to the material charge sensitivity coefficients and the crystal thickness, and inversely proportional to the crystal area and the material relative permittivity. The piezoelectric effect is reversible, so that a change in voltage also generates a force and a corresponding change in thickness. Thus the same device can be *both a sensor and an actuator*. Combined sensor/actuators are an intriguing topic of current research.

Tunneling Sensing depends on the exponential relationship between the tunneling current I and the tip/surface separation z given by $I = I_0 e^{-kz}$, where k depends on the tunnel barrier height in eV. Tunneling is an extremely accurate method of sensing nanometer-scale displacements, but its highly nonlinear nature requires the use of feedback control to make it useful.

Capacitive Sensors typically have one fixed plate and one movable plate. When a force is applied to the movable plate, the change in capacitance C is given as $\Delta C = \varepsilon A / \Delta d$, with Δd the resulting displacement, A the area, and ε the dielectric constant. Changes in capacitance can be detected using a variety of electric circuits and converted to a voltage or current change for further processing. **Inductive sensors**, which convert displacement to a change in inductance, are also often useful.

Magnetic and Electromagnetic Sensors do not require direct physical contact and are useful for detecting proximity effects [Kovacs 1998].

The Hall Effect, discovered by Edwin Hall in 1879, relies on the fact that the Lorentz Force deflects flowing charge carriers in a direction perpendicular to both their direction of flow and an applied magnetic field (i.e. vector cross product). The Hall voltage induced in a plate of thickness T is given by $V_H = RI_x B_z / T$, with R the Hall coefficient, I_x the current flow in direction x , and B_z the magnetic flux density in the z direction. R is 4-5 times larger in semiconductors than in most metals. The **Magnetoresistive effect** is a related phenomenon depending on the fact that the conductivity varies as the square of the applied flux density.

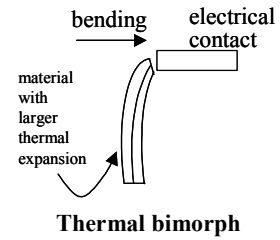


The Hall Effect

Magnetic Field Sensors can be used to detect the remote presence of metallic objects. **Eddy-Current Sensors** use magnetic probe coils to detect defects in metallic structures such as pipes.

Thermal Sensors are a family of sensors used to measure temperature or heat flux. Most biological organisms have developed sophisticated temperature sensing systems [Kovacs 1998].

Thermo-Mechanical Transduction is used for temperature sensing and regulation in homes and automobiles. On changes in temperature T , all materials exhibit (linear) thermal expansion of the form $\Delta L/L = \alpha \Delta T$, with L the length and α the coefficient of linear expansion. One can fabricate a strip of two joined materials with different thermal expansions. Then, the radius of curvature of this thermal bimorph depends on the temperature change.



Thermoresistive Effects are based on the fact that the resistance R changes with temperature T . For moderate changes, the relation is approximately given by for many metals by $\Delta R/R = \alpha_R \Delta T$, with α_R the temperature coefficient of resistance. The relationship for silicon is more complicated but is well understood. Hence, silicon is useful for detecting temperature changes.

Thermocouples are based on the thermoelectric *Seebeck effect*, whereby if a circuit consists of two different materials joined together at each end, with one junction hotter than the other, a current flows in the circuit. This generates a Seebeck voltage given approximately by $V \approx \alpha(T_1 - T_2) + \gamma(T_1^2 - T_2^2)$ with T_1, T_2 the temperatures at the two junctions. The coefficients depend on the properties of the two materials. Semiconductor thermocouples generally have higher sensitivities than do metal thermocouples. Thermocouples are inexpensive and reliable, and so are much used. Typical thermocouples have outputs on the order of $50 \mu\text{V}/^\circ\text{C}$ and some are effective for temperature ranges of -270°C to 2700°C .

Resonant Temperature Sensors rely on the fact that single-crystal SiO_2 exhibits a change in resonant frequency depending on temperature change. Since this is a frequency effect, it is more accurate than amplitude-change effects and has extreme sensitivity and accuracy for small temperature changes.

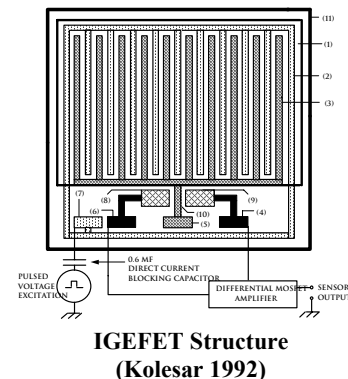
Optical Transducers convert light to various quantities that can be detected [Kovacs 1998]. These are based on one of several mechanisms. In the *photoelectric effect* (Einstein, Nobel Prize, 1921) one electron is emitted at the negative end of a pair of charged plates for each light photon of sufficient energy. This causes a current to flow. In *photoconductive sensors*, photons generate carriers that lower the resistance of the material. In *junction-based photosensors*, photons generate electron-hole pairs in a semiconductor junction that causes current flow. This is often misnamed the *photovoltaic effect*. These devices include photodiodes and phototransistors. *Thermopiles* use a thermocouple with one junction coated in a gold or bismuth black absorber, which generates heat on illumination.

Solar cells are large photodiodes that generate voltage from light. *Bolometers* consist of two thermally sensitive resistors in a Wheatstone bridge configuration, with one of them shielded from the incident light. Optical transducers can be optimized for different frequencies of light, resulting in infrared detectors, ultraviolet detectors, etc.

Various devices, including accelerometers, are based on *optical fiber technology*, often using time-of-flight information.

Chemical And Biological Transducers [Kovacs 1998] cover a very wide range of devices that interact with solids, liquids, and gases of all types. Potential applications include environmental monitoring, biochemical warfare monitoring, security area surveillance, medical diagnostics, implantable biosensors, and food monitoring. Effective use has been shown for NO_x (from pollution), organophosphorus pesticides, nerve gases (Sarin, etc), hydrogen cyanide, smallpox, anthrax, CO_x , SO_x , and others.

Chemiresistors have two interdigitated finger electrodes coated with specialized chemical coatings that change their resistance when exposed to certain chemical challenge agents. The electrodes may be connected directly to an FET, which amplifies the resulting signals *in situ* for good noise rejection. This device is known as an interdigitated-gate electrode FET (IGEFET). Arrays

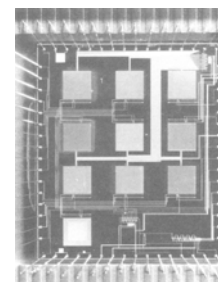


of chemiresistors, each device with a different chemically active coating, can be used to increase specificity for specific challenge agents [Kolesar 1992]. Digital signal processing, including neural network classification techniques, is important in correct identification of the agent.

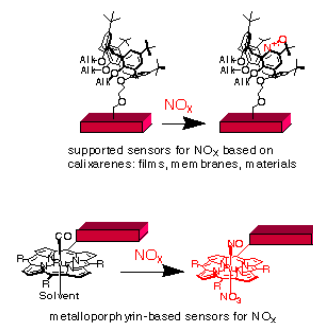
Metal-Oxide Gas Sensors rely on the fact that adsorption of gases onto certain semiconductors greatly changes their resistivities. In thin-film detectors, a catalyst such as platinum is deposited on the surface to speed the reactions and enhance the response. Useful as sensors are the oxides of tin, zinc, iron, zirconium, etc. Gases that can be detected include CO₂, CO, H₂S, NH₃, and ozone. Reactions are of the form $O_2 + 2e^- \rightarrow 2O^-$ so that adsorption effectively produces an electron trap site, effectively depleting the surface of mobile carriers and increasing its resistance.

Electrochemical Transducers rely on currents induced by oxidation or reduction of a chemical species at an electrode surface. These are among the simplest and most useful of chemical sensors. An electron transfer reaction occurs that is described by $O + ze^- \rightleftharpoons R$, with *O* the oxidized species, *R* the reduced species, and *z* the charge on the ion involved. The resulting current density is given in terms of *z* by the Butler-Volmer equation [Kovacs 1998].

Biosensors of a wide variety of types depend on the high selectivity of many biomolecular reactions, e.g. molecular binding sites of the detector may only admit certain species of analyte molecules. Unfortunately, such reactions are not usually reversible so the sensor is not reusable. These devices have a biochemically active thin film deposited on a platform device that converts induced property changes (e.g. mass, resistance) into detectable electric or optical signals. Suitable conversion platforms include the IGFET (above), ion-sensitive FET (ISFET), SAW (below), quartz crystal microbalance (QCM), microcantilevers, etc. To provide specificity to a prescribed analyte measurand, for the thin film one may use proteins (enzymes or antibodies), polysaccharide, nucleic acid, oligonucleotides [Choi, Gracy, et al. 2002], or an ionophore (which has selective responses to specific ion types). Arrays of sensors can be used, each having a different biochemically active film, to improve sensitivity. This has been used in the so-called ‘electronic nose.’

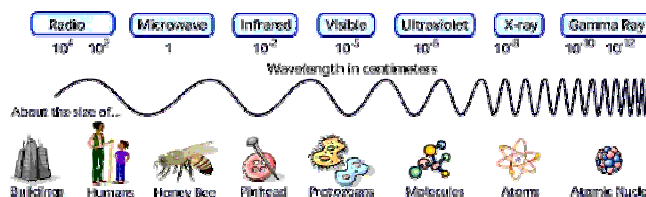


3x3 IGFET Sensor Microarray



Biosensors based on molecular recognition [Rudkevich 1996]

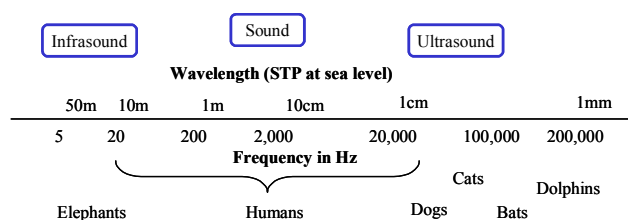
The Electromagnetic Spectrum can be used to fabricate *Remote Sensors* of a wide variety of types. Generally the wavelength suitable for a particular application is selected based on the propagation distance, the level of detail and resolution required, the ability to penetrate solid materials or certain mediums, and the signal processing difficulty. Doppler techniques allow the measurement of velocities. Millimeter waves have been used for satellite remote monitoring. Infrared is used for night vision and sensing heat. IR motion detectors are inexpensive and reliable. Electromagnetic waves can be used to determine distance using time-of-flight information- Radar uses RF waves and Lidar uses light (laser). The velocity of light is $c = 299.8 \times 10^6 \text{ m/s}$. GPS uses RF for absolute position localization. Visible light imaging using cameras is used in a broad range of applications but generally requires the use of sophisticated and computationally expensive DSP techniques including edge detection, thresholding, segmentation, pattern recognition, motion analysis, etc.



The electromagnetic spectrum

Courtesy of <http://imagers.gsfc.nasa.gov/ems/waves3.html>

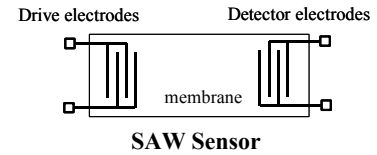
Acoustic Sensors include those that use sound as a sensing medium. Doppler techniques allow the measurement of velocities. Ultrasound often provides more information about mechanical machinery vibrations, fluid leakage, and impending equipment faults than do other techniques. Sonar uses sound to determine distance using time-of-flight information. It is effective in media other than air, including



The acoustic spectrum

underwater. Caution should be used in that the propagation speed of acoustic signals depends on the medium. The speed of sound at sea level in a standard atmosphere is $c_s=340.294$ m/s. Subterranean echoes from earthquakes and tremors can be used to glean information about the earth's core as well as about the tremor event, but deconvolution techniques must be used to remove echo phenomena and to compensate for uncertain propagation speeds.

Acoustic Wave Sensors are useful for a broad range of sensing devices [Kovacs 1998]. These transducers can be classified as surface acoustic wave (SAW), thickness-shear mode (TSM), flexural plate wave (FPW), or acoustic plate mode (APM). The SAW is shown in the figure and consists of two sets of interdigitated fingers at each end of a membrane, one set for generating the SAW and one for detecting it. Like the IGFET, these are useful platforms to convert property changes such as mass into detectable electrical signals. For instance, the surface of the device can be coated with a chemically or biologically active thin film. On presentation of the measurand to be sensed, adsorption might cause the mass m to change, resulting in a frequency shift given by the Sauerbrey equation $\Delta f = kf_o^2 \Delta m / A$, with f_o the membrane resonant frequency, constant k depending on the device, and A the membrane area.



2.3.3. Sensors for Smart Environments

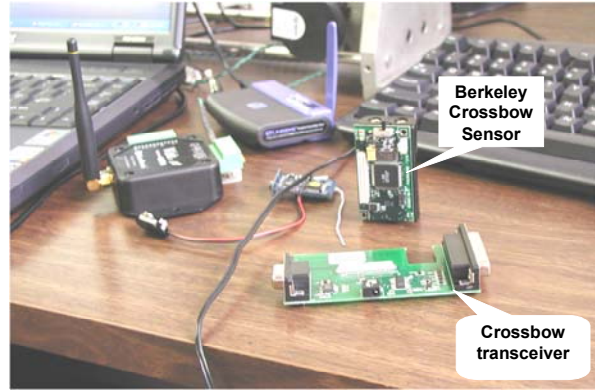
Many vendors now produce commercially available sensors of many types that are suitable for wireless network applications. See for instance the websites of SUNX Sensors, Schaevitz, Keyence, Turck, Pepperl & Fuchs, National Instruments, UE Systems (ultrasonic), Leake (IR), CSI (vibration). The table shows which physical principles may be used to measure various quantities. MEMS sensors are by now available for most of these measurands.

Measurements for Wireless Sensor Networks		
	Measurand	Transduction Principle
Physical Properties	Pressure	Piezoresistive, capacitive
	Temperature	Thermistor, thermo-mechanical, thermocouple
	Humidity	Resistive, capacitive
	Flow	Pressure change, thermistor
Motion Properties	Position	E-mag, GPS, contact sensor
	Velocity	Doppler, Hall effect, optoelectronic
	Angular velocity	Optical encoder
	Acceleration	Piezoresistive, piezoelectric, optical fiber
Contact Properties	Strain	Piezoresistive
	Force	Piezoelectric, piezoresistive
	Torque	Piezoresistive, optoelectronic
	Slip	Dual torque
	Vibration	Piezoresistive, piezoelectric, optical fiber, Sound, ultrasound
Presence	Tactile/contact	Contact switch, capacitive
	Proximity	Hall effect, capacitive, magnetic, seismic, acoustic, RF
	Distance/range	E-mag (sonar, radar, lidar), magnetic, tunneling
	Motion	E-mag, IR, acoustic, seismic (vibration)
Biochemical	Biochemical agents	Biochemical transduction
Identification	Personal features	Vision
	Personal ID	Fingerprints, retinal scan, voice, heat plume, vision motion analysis

2.3.4. Commercially Available Wireless Sensor Systems

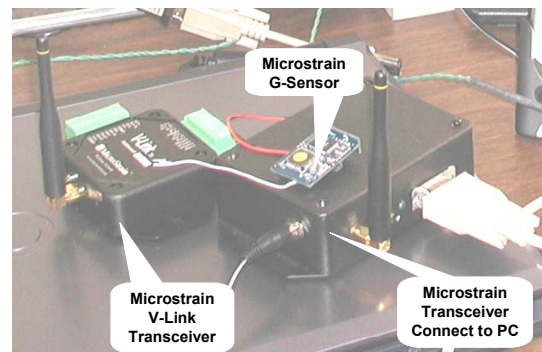
Many commercially available wireless communications nodes are available including Lynx Technologies, and various Bluetooth kits, including the Casira devices from Cambridge Silicon Radio, CSR.

Crossbow Berkeley Motes may be the most versatile wireless sensor network devices on the market for prototyping purposes. Crossbow (<http://www.xbow.com/>) makes three Mote processor radio module families– MICA [MPR300] (first generation), MICA2 [MPR400] and MICA2-DOT [MPR500] (second generation). Nodes come with five sensors installed- Temperature, Light, Acoustic (Microphone), Acceleration/Seismic, and Magnetic. These are especially suitable for surveillance networks for personnel and vehicles. Different sensors can be installed if desired. Low power and small physical size enable placement virtually anywhere. Since all sensor nodes in a network can act as base stations, the network can self configure and has multi-hop routing capabilities. The operating frequency is ISM band, either 916Mhz or 433 MHz, with a data rate of 40 Kbits/sec. and a range of 30 ft to 100 ft. Each node has a low power microcontroller processor with speed of 4MHz, a flash memory with 128 Kbytes, and SRAM and EEPROM of 4K bytes each. The operating system is Tiny-OS, a tiny micro-threading distributed operating system developed by UC Berkeley, with a NES-C (Nested C) source code language (similar to C). Installation of these devices requires a great deal of programming. A workshop is offered for training.



Berkeley Crossbow Motes

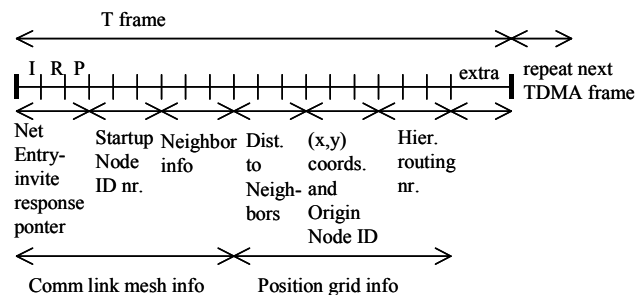
Microstrain’s X-Link Measurement System (<http://www.microstrain.com/>) may be the easiest system to get up and running and to program. The frequency used is 916 MHz, which lies in the US license-free ISM band. The sensor nodes are multi-channel, with a maximum of 8 sensors supported by a single wireless node. There are three types of sensor nodes – S-link (strain gauge), G-link (accelerometer), and V-link (supports any sensors generating voltage differences). The sensor nodes have a pre-programmed EPROM, so a great deal of programming by the user is not needed. Onboard data storage is 2MB. Sensor nodes use a 3.6-volt lithium ion internal battery (9V rechargeable external battery is supported). A single receiver (Base Station) addresses multiple nodes. Each node has a unique 16-bit address, so a maximum of 2^{16} nodes can be addressed. The RF link between Base Station and nodes is bi-directional and the sensor nodes have a programmable data logging sample rate. The RF link has a 30 meter range with a 19200 baud rate. The baud rate on the serial RS-232 link between the Base Station and a terminal PC is 38400. LabVIEW interface is supported.



Microstrain Wireless Sensors

2.3.5. Self-Organization and Localization

Ad hoc networks of nodes may be deployed using, e.g. aircraft or ships. Self organization of ad hoc networks includes both *communications* self-organization and *positioning* self-organization. In the former, the nodes must wake up, detect each other, and form a communication network. Technologies for this are by now standard, by and large developed within the mobile phone industry. Distributed surveillance sensor networks require information about the relative positions of the nodes for distributed signal processing, as well as absolute positioning information for reporting data related to detected targets.



TDMA frame for communication protocols and localization

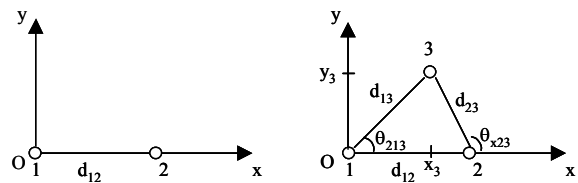
Relative Layout Positioning- Localization. Relative positioning or *localization* requires internode communications, and a TDMA message header frame that has both communications and localization fields is shown in the figure. There are various means for a node to measure distance to its neighbors, mostly based on RF time-of-flight information. In air, the propagation speed is known, so time differences can be converted to distances. Given the relative distances between nodes, we want to organize the web into a grid specified in terms of relative positions.

An approach based on *robot kinematic transformations* provides a straightforward iterative technique for adding new nodes to a network. A homogeneous transformation is a 4x4 matrix [Lewis, Abdallah, Dawson 1993]

$$A_i = \begin{bmatrix} R_i & p_i \\ 0 & 1 \end{bmatrix}$$

where R_i is a 3x3 rotation matrix and $p_i \in R^3$ is a translation vector. The T matrix defined by iterative rotations and translations as $T_{ij} = A_i A_{i+1} \dots A_j$ allows one to express vectors in frame j in terms of the coordinates of frame i . If the network is a flat 2D net, the z coordinates can be ignored, simplifying the problem.

The figure shows how to start a self-organizing algorithm for relative positioning location. The first node to wake up is assigned the origin O . As nodes wake up they are invited into the grid and distance is determined. The second node, a distance d_{12} from the first, defines the x and y axes. When the third node is discovered and two distances measured, one computes its x and y coordinates as follows. Let A_{ij} denote the A matrix relating points i and j . Then, in standard fashion (c.f. 2-link robot arm) [Lewis et al. 1993] one computes A_{12} , A_{13} , A_{23} . in terms of d_{12} , d_{13} , d_{23} . One can write the relative location in frame O of the new point 3 in two ways: $T_{13} = A_{13}$ and $T_{123} = A_{12}A_{23}$. The triangle shown in the figure is a *closed kinematic chain* of the sort studied in [Liu and Lewis 1993]. The solution is obtained by requiring that the two maps T_{13} and T_{123} be *exact* at point 3. This means that the position vectors p_{13} and p_{123} (i.e. the third columns) of the two maps must be the same. This results in a nonlinear equation that can be solved for the distances.



a. Two nodes- define x & y axes

b. 3 node closed kinematic chain- compute (x_3, y_3)

Integrating new nodes into relative positioning grid

Homogeneous transformations allow for a *fast recursive procedure for integrating new nodes* into the grid. Suppose that a new node, number 4, enters the net. For unique positioning it must find distances to *three nodes* already in the grid. Then, based on relative distance information, one computes A matrices and T matrices to interrelate nodes 1,2,3,4. Then, x and y coordinates of node 4 relative to the origin are computed uniquely by forcing three maps to be exact at node 4. That is, $A_{14} = A_{12}A_{24} = A_{13}A_{34}$. Now, the coordinates of the new node in terms of the base frame O for the subgrid can be computed.

Absolute Geographical Positioning. A network is said to be *relatively calibrated* if the relative positions of all nodes are known. Now, it is necessary to determine the absolute geographic position of the network. The net is said to be (fully) *calibrated* if the absolute positions of all nodes are known. To determine the absolute node positions in a relatively calibrated flat 2D net, at least three nodes in the net must determine their absolute positions. There are many ways for a node to determine its absolute position, including GPS and techniques based on stored maps, landmarks, or beacons [Bulusu and Estrin 2002].

Ultra Wideband Radio. UWB is of great interest recently for communications in distributed sensor networks. This is because UWB is a short-range technology that can penetrate walls, it is suitable for multi-node transmissions, and it has built in time-of-flight properties that make it very easy to measure ranges down to 1 cm with a range of 40m. This means that the *same medium*, UWB, can be used for communications, localization, and target tracking in a distributed surveillance network. Moreover, UWB transceivers can be made very small and are amenable for MEMS technology; since pulse position modulation (PPM) is used, no carrier is needed, meaning that antennas are not inductive. Also, the receiver is based on a rake detector and correlator bank so that no IF stage is needed.

UWB uses signals like [Ray 2001]

$$s(t) = \sum_j w(t - jT_f - c_j T_c - \delta d_{\lfloor j/N_s \rfloor})$$

where $w(t)$ is the basic pulse of duration approx. 1ns, often a wavelet or a Gaussian monocycle, and T_f is the frame or pulse repetition time. In a multi-node environment, catastrophic collisions are avoided by using a pseudorandom sequence c_j to shift pulses within the frame to different compartments, and the compartment size is T_c sec. One may have, for instance, $T_f = 1\mu\text{sec}$ and $T_c = 5\text{ns}$. Data is transmitted using digital PPM, where if the data bit is 0 the pulse

is not shifted, and if the data bit is 1 the pulse is shifted by δ . The modulation shift is selected to make the correlation of $w(t)$ and $w(t-\delta)$ as negative as possible. The meaning of $d_{\lfloor j/N_s \rfloor}$ is that the same data bit is transmitted N_s times, allowing for very reliable communications with low probability of error.

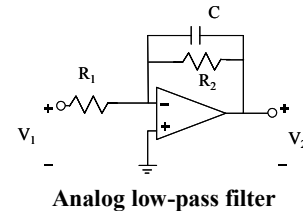
2.4. SIGNAL PROCESSING AND DECISION-MAKING

The figure showing the IEEE 1451 Smart Sensor includes basic blocks for signal conditioning (SC), digital signal processing (DSP), and A/D conversion. Let us briefly mention some of the issues here.

2.4.1. Signal Conditioning

Signals coming from MEMS sensors can be very noisy, of low amplitude, biased, and dependent on secondary parameters such as temperature. Moreover, one may not always be able to measure the quantity of interest, but only a related quantity. Therefore signal conditioning is usually required. SC is performed using electronic circuitry, which may conveniently be built using standard VLSI fabrication techniques *in situ* with MEMS sensors. A reference for SC, A/D conversion, and filtering is [Lewis 1992].

A real problem with MEMS sensors is undesired sensitivity to secondary quantities such as temperature. *Temperature compensation* can often be directly built into a MEMS sensor circuit. In the figure above showing a 3x3 array of IGFET sensors, there is shown a 10th IGFET- this is for temperature compensation. Temperature compensation can also be added during the SC stage as discussed below.



A basic technique for improving the signal-to-noise ratio (SNR) is *low-pass filtering*, since noise generally dominates the desirable signals at high frequencies. Shown in the figure is an analog LPF that also amplifies, constructed from an operational amplifier. Such devices are easily fabricated using VLSI semiconductor techniques. The time constant of this circuit is $\tau = R_2 C$. The transfer function of this filter is $H(s) = k a / (s + a)$ with 3 dB cutoff frequency given by $a = 1/\tau$ rad. and gain given by $k = R_2 / R_1$. Here, s is the Laplace transform variable. The cutoff frequency should be chosen larger than the highest useful signal frequency of the sensor.

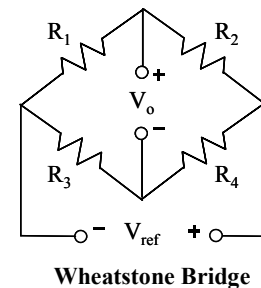
Alternatively, one may use a digital LPF implemented on a computer after sampling. A digital low-pass filter transfer function and the associated difference equation for implementation is given by

$$\text{Digital filter: } \hat{s}_k = K \frac{z+1}{z-\alpha} s_k \qquad \text{difference equation: } \hat{s}_{k+1} = \alpha s_k + K(s_{k+1} + s_k)$$

Here, z is the z-transform variable treated as a unit delay in the time domain, s_k is the measured signal, and \hat{s}_k is the *filtered* or *smoothed variable* with reduced noise content. The filter parameters are selected in terms of the desired cutoff frequency and the sampling period [Lewis 1992].

It is often the case that one can measure a variable s_k (e.g. position), but needs to know its rate of change v_k (e.g. velocity). Due to the presence of noise, one cannot simply take the difference between successive values of s_k as the velocity. A filtered velocity estimate given by $v_{k+1} = \alpha v_k + K(s_{k+1} - s_k)$ both filters out noise and gives a smooth velocity estimate.

Often, changes in resistance must be converted to voltages for further processing. This may be accomplished by using a Wheatstone bridge [de Silva 1989]. Suppose $R_1 = R$ in the figure is the resistance that changes depending on the measurand (e.g. strain gauge), and the other three resistances are constant (quarter bridge configuration). Then the output voltage changes according to $\Delta V_o = V_{ref} \Delta R / 4R$. We assume a balanced bridge so that $R_2 = R_1 = R$ and $R_3 = R_4$. Sensitivity can be improved by having two sensors *in situ*, such that the changes in each are opposite (e.g. two strain gauges on opposite sides of a flexing bar). This is known as a half bridge. If R_1 and R_2 are two such sensors and $\Delta R_1 = -\Delta R_2$, then the output voltage doubles. The Wheatstone bridge may also be used for differential measurements (e.g. for insensitivity to common changes of two sensors), to improve sensitivity, to remove zero offsets, for temperature compensation, and to perform other signal conditioning.



Specially designed operational amplifier circuits are useful for general signal conditioning [Frank 2000]. Instrumentation Amplifiers provide differential input and common mode rejection, impedance matching between

sensors and processing devices, calibration, etc. SLEEPMODE amplifiers (Semiconductor Components Ind., LLC) consume minimum power while asleep, and activate automatically when the sensor signal exceeds a prescribed threshold.

2.4.2. Digital Signal Processing

Sensor fusion is important in a network of sensors of different modalities. A distributed vehicle/personnel surveillance network might include seismic, acoustic, infrared motion, temperature, and magnetic sensors. The standard DSP tool for combining the information from many sensors is the *Kalman Filter* [Lewis 1986, 1992]. The Kalman Filter is used for communications, navigation, feedback control, and elsewhere and provides the accuracy that allowed man to navigate in space and eventually to reach the moon and more recently to send probes to the limits of the Solar System.

A properly designed Kalman Filter allows one to observe only a few quantities, or *measured outputs*, and then reconstruct or estimate the full internal state of a system. It also provides low-pass filtering functions and amplification, and can be constructed to provide temperature compensation, common mode rejection, zero offset correction, etc. The *discrete-time Kalman Filter*, useful for DSP using microprocessors, is a dynamical filter given by

$$\hat{x}_{k+1} = A(I - KH)\hat{x}_k + Bu_k + AKz_k$$

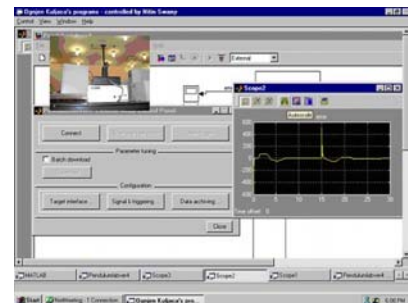
where the sensed outputs are in a vector z_k , the control inputs to the system being observed are in vector u_k , and the estimates of the internal states are given by the vector \hat{x}_k . Note that the number of sensed outputs can be significantly less than the number of states one can estimate. In this filter, matrices A and B represent the known dynamics of the sensed system, and the sensed outputs are given as a linear combination of the states by $z_k = Hx_k$, where H is a known measurement matrix. The Kalman gain K is determined by solving a design equation known as the Riccati Equation. The Kalman Filter is the optimal linear estimator given the known system properties and prescribed corrupting noise statistics.

Distributed signal processing is the most efficient means of computation in a network of distributed signal-processing nodes. The theory of *Decentralized Kalman Filtering* provides a formal mechanism for apportioning sensor filtering, reconstruction, and compensation tasks among a hierarchically organized group of nodes.

Other DSP tools include techniques used in spectrum analysis, speech processing, stock market analysis, etc. Statistical methods allow regression analysis, correlation analysis, principal component analysis, and clustering. Also available are a wide range of techniques based on neural network properties of classification, association, generalization, and clustering. Decision-making paradigms include fuzzy logic, Bayesian decision-making, Dempster-Shafer, diagnostic/prescription-based schemes as used in the medical field, and so on. The MathWorks software MATLAB has extensive capabilities in all these areas, and specialized Toolboxes provide powerful tools for DSP and decision-making for distributed wireless sensor networks.

2.4.3. Decision-Making and User Interface

Many software products are available to provide advanced DSP, intelligent user interfaces, decision assistance, and alarm functions. Among the most popular, powerful, and easy to use is National Instruments LabVIEW software. Available are toolkits for camera image processing, machinery signal processing and diagnostics, sensor equipment calibration and test, feedback control, and more. The figure shows a LabVIEW user interface for monitoring machinery conditions over the Internet for automated maintenance scheduling functions. Included are displays of sensor signals that can be selected and tailored by the user. The user can prescribe bands of normal operation, excursions outside of which generate alarms of various sorts and severity.



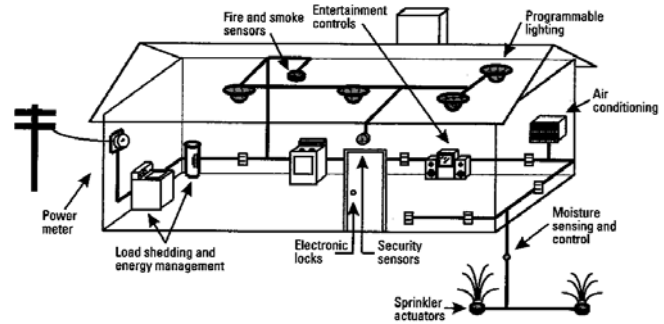
LabVIEW user interface for wireless Internet-based remote site monitoring

2.5. Building and Home Automation

The figure shows how networks of various sorts might interact in the smart home environment. An excellent reference for this section is Frank [2000]. There are many available protocols for networking of the smart home, and it is not necessary to develop new protocols on one's own for commercially acceptable systems. The BACnet protocol has been developed by the building automation industry to provide a standard for interconnecting networks for building sensing and control. Networks that can be used include Ethernet, MS/TP, and LonWorks. Building energy management standards are being developed by the American Society of Heating, Refrigeration, and Air-

Conditioning Engineers (ASHRAE). A major driver for the smart home is the power distribution industry, which could save enormous sums with demand-side regulation and automated remote meter reading. The Intelligent Building Institute has been a force in developing appropriate standards.

The **X-10 protocol** is used for lamp and appliance controls. The more recently developed **Smart House Applications Language (SHAL)** includes over 100 message types for specific sensing and control functions. However, SHAL requires dedicated multiconductor wiring. The **Consumer Electronics bus (CEBus)**, initiated by the Electronic Industries Association, provides both data and control channels and handles up to 10Kbps. It is useful for the utility industry.



Smart Home Networks
Courtesy of Artech House, R. Frank 2000

Several automotive protocols have been developed, and some of these are useful also for building control. **CAN** is a serial communications protocol developed for automotive multiplex wiring systems, and has been adopted in industrial applications by manufacturers such as Allen-Bradley (in the DeviceNET system) and Honeywell (in SDS). CAN supports distributed real-time control with a high level of security, and is a multimaster protocol that allows any node in the network to communicate with any other node. Supported are user-defined message prioritization, multiple access/collision resolution, and error detection.

The **LonWorks protocol**, developed by Echelon Corp (<http://www.echelon.com/products/lonworks/default.htm>) is very convenient for industrial and consumer applications. It supports all seven layers of the OSI/RM model, and supports fieldbus requirements, arbitration, and message coding. LonWorks operates on a peer-to-peer bus network basis. Devices in a LonWorks network communicate using LonTalk. This language provides a set of services that allow the application program in a device to send and receive messages from other devices over the network without needing to know the topology of the network or the names, addresses, or functions of other devices. The LonWorks protocol can optionally provide end-to-end acknowledgement of messages, authentication of messages, and priority delivery to provide bounded transaction times. Support for network management services allows for remote network management tools to interact with devices over the network, including reconfiguration of network addresses and parameters, downloading of application programs, reporting of network problems, and start/stop/reset of device application programs. LonWorks networks can be implemented over basically any medium, including power lines, twisted pair, radio frequency (RF), infrared (IR), coaxial cable and fiber optics.

REFERENCES

- E. Altman, T. Basar, T. Jimenez, and N. Shimkin, "Competitive routing in networks with polynomial costs," *IEEE Trans. Automat. Control*, vol. 47, no. 1, pp. 92-96, 2002.
- R. Bronson and G. Naadimuthu, *Operations Research*, 2 ed., Schaum's Outlines, McGraw Hill, New York, 1997.
- N. Bulusu, J. Heidemann, D. Estrin, and T. Tran, "Self-configuring localization systems: design and experimental evaluation," pp. 1-31, *ACM TECS special Issue on Networked Embedded Computing*, Aug. 2002.
- J. Cao and F. Zhang, "Optimal configuration in hierarchical network routing," *Proc. Canadian Conf. Elect. and Comp. Eng.*, pp. 249-254, Canada 1999.
- T.-S. Chen, C.-Y. Chang, and J.-P. Sheu, "Efficient path-based multicast in wormhole-routed mesh networks," *J. Sys. Architecture*, vol. 46, pp. 919-930, 2000.
- J. Choi, C. Conrad, C. Malakowsky, J. Talent, C.S. Yuan, and R.W. Gracy, "Flavones from *Scutellaria baicalensis* Georgi attenuate apoptosis and protein oxidation in neuronal cell lines," *Biochemica et Biophysica Acta*, 1571: 201-210 (2002).
- Conway and Heffernan, Univ. Limerick, 2003, <http://www.ul.ie/~pei>
- C.W. de Silva, *Control Sensors and Actuators*, Prentice-Hall, New Jersey, 1989.
- J. Duato, "A necessary and sufficient condition for deadlock-free routing in cut-through and store-and-forward networks," *IEEE Trans Parallel and Distrib. Systems*, vol. 7, no. 8, pp. 841-854, Aug. 1996.
- R. Frank, *Understanding Smart Sensors*, 2nd Ed., Artech House, Norwood, MA, 2000.
- M.R. Garey, and D.S. Johnson, *Computers and Intractability: a Guide to the Theory of NP-completeness*. Freeman, San Francisco, CA, 1979.
- F. Giulietti, L. Pollini, and M. Innocenti, "Autonomous formation flight," *IEEE Control Systems Mag.*, pp. 34-44, Dec. 2000.

IEEE 1451, A Standard Smart Transducer Interface, Sensors Expo, Philadelphia, Oct. 2001, http://iee1451.nist.gov/Workshop_04Oct01/1451_overview.pdf

R. Jordan and C.A. Abdallah, "Wireless communications and networking: an overview," Report, Elect. and Comp. Eng. Dept., Univ. New Mexico, 2002.

E.S. Kolesar, C.P. Brothers, C.P. Howe, et al., "Integrated circuit microsensor for selectively detecting nitrogen dioxide and diisopropyl methylphosphate," *Thin Solid Films*, vol. 220, pp. 30-37. 1992.

G.T.A. Kovacs, *Micromachined Transducers Sourcebook*, McGraw-Hill, Boston, 1998.

P.R. Kumar, "New technological vistas for systems and control: the example of wireless networks," *IEEE Control Systems Magazine*, pp. 24-37, Feb. 2001.

F.L. Lewis, *Optimal Estimation*, Wiley, New York, 1986.

F.L. Lewis, *Applied Optimal Control and Estimation*, Prentice-Hall, New Jersey, 1992.

F.L. Lewis, C.T. Abdallah, and D.M Dawson, *Control of Robot Manipulators*, Macmillan, New York, Mar. 1993.

F.L. Lewis and V.L. Syrmos, *Optimal Control*, 2nd ed., Wiley, New York, 1995.

K. Liu, M. Fitzgerald, and F.L. Lewis, "Kinematic analysis of a Stewart Platform manipulator," *IEEE Trans. Industrial Electronics*, vol. 40, no. 2, pp. 282-293, 1993.

S.H. Low, F. Paganini, and J.C. Doyle, "Internet congestion control," *IEEE Control Systems Mag.*, pp. 28-43, Feb. 2002.

M. Madou, *Fundamentals of Microfabrication*, CRC Press, Boca Raton, 1997.

PC Tech Guide, 2003, <http://pctechguide.com/29network.htm>

S. Ray, "An introduction to ultra wide band (impulse) radio," Internal Report, Elect. and Computer Eng. Dept, Boston Univ., Oct. 2001.

D.M. Rudkevich, J. Scheerder, and D.N. Reinhoudt, "Anion recognition by natural receptors," in *Molecular Design and Bioorganic Catalysis*, ed. C.S. Wilcox, pp. 137-162, Kluwer, Boston, 1996.

S. Shah-Heydari and O. Yang, "A tree-based algorithm for protection/restoration in optical mesh networks," *Proc. Canadian Conf. Elect. and Comp. Eng.*, vol. 2, pp. 1169-1174, Canada 2001.

J. Shi and J.P. Fonseka, "Hierarchical self-healing rings," *IEEE/ACM Trans. Networking*, vol. 3, no. 6, pp. 690-697, Dec. 1995.

Smart Transducer Interface Standard, IEEE 1451, *Sensors Expo*, Philadelphia, Oct. 2001.

J.W. Smith, Rand Corp., *On Distributed Communications*, Memorandum RM-3578-PR, 1964, <http://www.rand.org/publications/RM/RM3578/>

C. Smythe, "ISO 8802/5 token ring local area networks," *Elect. & Communic. Eng. J.*, vol. 11, no. 4, pp. 195-207, Aug. 1999.

N. Swamy, *Control Algorithms for Networked Control and Communication Systems*, PhD Thesis, Dept. of Elect. Eng., The University of Texas at Arlington, Texas, 2003.