1

Introduction

Objectives of this Chapter

Applications should shape and form the technology for which they are intended. This holds true in particular for wireless sensor networks, which have, to some degree, been a technology-driven development. This chapter starts out by putting the idea of wireless sensor networks into a broader perspective and gives a number of application scenarios, which will later be used to motivate particular technical needs. It also generalizes from specific examples to types or classes of applications. Then, the specific challenges for these application types are discussed and why current technology is not up to meeting these challenges.

At the end of this chapter, the reader should have an appreciation for the types of applications for which wireless sensor networks are intended and a first intuition about the types of technical solutions that are required, both in hardware and in networking technologies.

Chapter Outline

1.1 The vision of Ambient Intelligence

The most common form of information processing has happened on large, general-purpose computational devices, ranging from old-fashioned mainframes to modern laptops or palmtops. In many applications, like office applications, these computational devices are mostly used to process information that is at its core centered around a human user of a system, but is at best indirectly related to the physical environment.

Protocols and Architectures for Wireless Sensor Networks H. Karl and A. Willig

In another class of applications, the physical environment is at the focus of attention. Computation is used to exert control over physical processes, for example, when controlling chemical processes in a factory for correct temperature and pressure. Here, the computation is integrated with the control; it is *embedded* into a physical system. Unlike the former class of systems, such **embedded systems** are usually not based on human interaction but are rather required to work without it; they are intimately tied to their control task in the context of a larger system.

Such embedded systems are a well-known and long-used concept in the engineering sciences (in fact, estimates say that up to 98 % of all computing devices are used in an embedded context [91]). Their impact on everyday life is also continuing to grow at a quick pace. Rare is the household where embedded computation is not present to control a washing machine, a video player, or a cell phone. In such applications, embedded systems meet human-interaction-based systems.

Technological progress is about to take this spreading of embedded control in our daily lives a step further. There is a tendency not only to equip larger objects like a washing machine with embedded computation and control, but also smaller, even dispensable goods like groceries; in addition, living and working spaces themselves can be endowed with such capabilities. Eventually, computation will surround us in our daily lives, realizing a vision of "**Ambient Intelligence**" where many different devices will gather and process information from many different sources to both control physical processes and to interact with human users. These technologies should be unobtrusive and be taken for granted – Marc Weiser, rightfully called the *father of ubiquitous computing*, called them *disappearing technologies* [867, 868]. By integrating computation and control in our physical environment, the well-known interaction paradigms of person-to-person, person-to-machine and machine-to-machine can be supplemented, in the end, by a notion of person-to-physical world [783]; the interaction with the physical world becomes more important than mere symbolic data manipulation [126].

To realize this vision, a crucial aspect is needed in addition to computation and control: communication. All these sources of information have to be able to transfer the information to the place where it is needed – an actuator or a user – and they should collaborate in providing as precise a picture of the real world as is required. For some application scenarios, such networks of sensors and actuators are easily built using existing, wired networking technologies. For many other application types, however, the need to wire together all these entities constitutes a considerable obstacle to success: Wiring is expensive (figures of up to US$200 per sensor can be found in the literature [667]), in particular, given the large number of devices that is imaginable in our environment; wires constitute a maintenance problem; wires prevent entities from being mobile; and wires can prevent sensors or actuators from being close to the phenomenon that they are supposed to control. Hence, *wireless communication* between such devices is, in many application scenarios, an inevitable requirement.

Therefore, a new class of networks has appeared in the last few years: the so-called Wireless Sensor Network (WSN) (see e.g. [17, 648]). These networks consist of individual nodes that are able to interact with their environment by sensing or controlling physical parameters; these nodes have to collaborate to fulfill their tasks as, usually, a single node is incapable of doing so; and they use wireless communication to enable this collaboration. In essence, the nodes without such a network contain at least some computation, wireless communication, and sensing or control functionalities. Despite the fact that these networks also often include actuators, the term wireless sensor network has become the commonly accepted name. Sometimes, other names like "wireless sensor and actuator networks" are also found.

These WSNs are powerful in that they are amenable to support a lot of very different real-world applications; they are also a challenging research and engineering problem because of this very flexibility. Accordingly, there is no single set of requirements that clearly classifies all WSNs, and there is also not a single technical solution that encompasses the entire design space. For example, in many WSN applications, individual nodes in the network cannot easily be connected to a wired power supply but rather have to rely on onboard batteries. In such an application, the energy efficiency of any proposed solution is hence a very important figure of merit as a long operation time is usually desirable. In other applications, power supply might not be an issue and hence other metrics, for example, the accuracy of the delivered results, can become more important. Also, the acceptable size and costs of an individual node can be relevant in many applications. Closely tied to the size is often the capacity of an onboard battery; the price often has a direct bearing on the quality of the node's sensors, influencing the accuracy of the result that can be obtained from a single node. Moreover, the number, price, and potentially low accuracy of individual nodes is relevant when comparing a distributed system of many sensor nodes to a more centralized version with fewer, more expensive nodes of higher accuracy. Simpler but numerous sensors that are close to the phenomenon under study can make the architecture of a system both simpler and more energy efficient as they facilitate distributed sampling – detecting objects, for example, requires a distributed system [17, 648].

Realizing such wireless sensor networks is a crucial step toward a deeply penetrating Ambient Intelligence concept as they provide, figuratively, the "last 100 meters" of **pervasive control**. To realize them, a better understanding of their potential applications and the ensuing requirements is necessary, as is an idea of the enabling technologies. These questions are answered in the following sections; a juxtaposition of wireless sensor networks and related networking concepts such as fieldbuses or mobile ad hoc network is provided as well.

1.2 Application examples

The claim of wireless sensor network proponents is that this technological vision will facilitate many existing application areas and bring into existence entirely new ones. This claim depends on many factors, but a couple of the envisioned application scenarios shall be highlighted.

Apart from the need to build cheap, simple to program and network, potentially long-lasting sensor nodes, a crucial and primary ingredient for developing actual applications is the actual sensing and actuating faculties with which a sensor node can be endowed. For many physical parameters, appropriate sensor technology exists that can be integrated in a node of a WSN. Some of the few popular ones are temperature, humidity, visual and infrared light (from simple luminance to cameras), acoustic, vibration (e.g. for detecting seismic disturbances), pressure, chemical sensors (for gases of different types or to judge soil composition), mechanical stress, magnetic sensors (to detect passing vehicles), potentially even radar (see references [245, 246] for examples). But even more sophisticated sensing capabilities are conceivable, for example, toys in a kindergarten might have tactile or motion sensors or be able to determine their own speed or location [783].

Actuators controlled by a node of a wireless sensor network are perhaps not quite as multifaceted. Typically, they control a mechanical device like a servo drive, or they might switch some electrical appliance by means of an electrical relay, like a lamp, a bullhorn, or a similar device.

On the basis of nodes that have such sensing and/or actuation faculties, in combination with computation and communication abilities, many different kinds of applications can be constructed, with very different types of nodes, even of different kinds within one application. A brief list of scenarios should make the vast design space and the very different requirements of various applications evident. Overviews of these and other applications are included in references [17, 26, 88, 91, 110, 126, 134, 245, 246, 351, 367, 392, 534, 648, 667, 783, 788, 803, 923].

Disaster relief applications One of the most often mentioned application types for WSN are disaster relief operations. A typical scenario is wildfire detection: Sensor nodes are equipped with thermometers and can determine their own location (relative to each other or in absolute coordinates). These sensors are deployed over a wildfire, for example, a forest, from an airplane. They collectively produce a "temperature map" of the area or determine the perimeter of areas with high temperature that can be accessed from the outside, for example, by firefighters equipped with Personal Digital Assistants (PDAs). Similar scenarios are possible for the control of accidents in chemical factories, for example.

Some of these disaster relief applications have commonalities with military applications, where sensors should detect, for example, enemy troops rather than wildfires. In such an application, sensors should be cheap enough to be considered disposable since a large number is necessary; lifetime requirements are not particularly high.

Environment control and biodiversity mapping WSNs can be used to control the environment, for example, with respect to chemical pollutants – a possible application is garbage dump sites. Another example is the surveillance of the marine ground floor; an understanding of its erosion processes is important for the construction of offshore wind farms. Closely related to environmental control is the use of WSNs to gain an understanding of the number of plant and animal species that live in a given habitat (biodiversity mapping).

The main advantages of WSNs here are the long-term, unattended, wirefree operation of sensors close to the objects that have to be observed; since sensors can be made small enough to be unobtrusive, they only negligibly disturb the observed animals and plants. Often, a large number of sensors is required with rather high requirements regarding lifetime.

Intelligent buildings Buildings waste vast amounts of energy by inefficient Humidity, Ventilation, Air Conditioning (HVAC) usage. A better, real-time, high-resolution monitoring of temperature, airflow, humidity, and other physical parameters in a building by means of a WSN can considerably increase the comfort level of inhabitants and reduce the energy consumption (potential savings of two quadrillion British Thermal Units in the US alone have been speculated about [667]). Improved energy efficiency as well as improved convenience are some goals of "intelligent buildings" [415], for which currently wired systems like BACnet, LonWorks, or KNX are under development or are already deployed [776]; these standards also include the development of wireless components or have already incorporated them in the standard.

In addition, such sensor nodes can be used to monitor mechanical stress levels of buildings in seismically active zones. By measuring mechanical parameters like the bending load of girders, it is possible to quickly ascertain via a WSN whether it is still safe to enter a given building after an earthquake or whether the building is on the brink of collapse – a considerable advantage for rescue personnel. Similar systems can be applied to bridges. Other types of sensors might be geared toward detecting people enclosed in a collapsed building and communicating such information to a rescue team.

The main advantage here is the collaborative mapping of physical parameters. Depending on the particular application, sensors can be retrofitted into existing buildings (for HVAC-type applications) or have to be incorporated into the building already under construction. If power supply is not available, lifetime requirements can be very high – up to several dozens of years – but the number of required nodes, and hence the cost, is relatively modest, given the costs of an entire building.

Facility management In the management of facilities larger than a single building, WSNs also have a wide range of possible applications. Simple examples include keyless entry applications where people wear badges that allow a WSN to check which person is allowed to enter which areas of a larger company site. This example can be extended to the detection of intruders, for example of vehicles that pass a street outside of normal business hours. A wide-area WSN could track such a vehicle's position and alert security personnel – this application shares many commonalities with corresponding military applications. Along another line, a WSN could be used in a chemical plant to scan for leaking chemicals.

These applications combine challenging requirements as the required number of sensors can be large, they have to collaborate (e.g. in the tracking example), and they should be able to operate a long time on batteries.

Machine surveillance and preventive maintenance One idea is to fix sensor nodes to difficult-to-reach areas of machinery where they can detect vibration patterns that indicate the need for maintenance. Examples for such machinery could be robotics or the axles of trains. Other applications in manufacturing are easily conceivable.

The main advantage of WSNs here is the cablefree operation, avoiding a maintenance problem in itself and allowing a cheap, often retrofitted installation of such sensors. Wired power supply may or may not be available depending on the scenario; if it is not available, sensors should last a long time on a finite supply of energy since exchanging batteries is usually impractical and costly. On the other hand, the size of nodes is often not a crucial issue, nor is the price very heavily constrained.

Precision agriculture Applying WSN to agriculture allows precise irrigation and fertilizing by placing humidity/soil composition sensors into the fields. A relatively small number is claimed to be sufficient, about one sensor per 100 m × 100 m area. Similarly, pest control can profit from a high-resolution surveillance of farm land. Also, livestock breeding can benefit from attaching a sensor to each pig or cow, which controls the health status of the animal (by checking body temperature, step counting, or similar means) and raises alarms if given thresholds are exceeded.

Medicine and health care Along somewhat similar lines, the use of WSN in health care applications is a potentially very beneficial, but also ethically controversial, application. Possibilities range from postoperative and intensive care, where sensors are directly attached to patients – the advantage of doing away with cables is considerable here – to the long-term surveillance of (typically elderly) patients and to automatic drug administration (embedding sensors into drug packaging, raising alarms when applied to the wrong patient, is conceivable). Also, patient and doctor tracking systems within hospitals can be literally life saving.

Logistics In several different logistics applications, it is conceivable to equip goods (individual parcels, for example) with simple sensors that allow a simple tracking of these objects during transportation or facilitate inventory tracking in stores or warehouses.

In these applications, there is often no need for a sensor node to *actively* communicate; passive readout of data is often sufficient, for example, when a suitcase is moved around on conveyor belts in an airport and passes certain checkpoints. Such passive readout is much simpler and cheaper than the active communication and information processing concept discussed in the other examples; it is realized by so-called Radio Frequency Identifier (RF ID) tags.

On the other hand, a simple RFID tag cannot support more advanced applications. It is very difficult to imagine how a passive system can be used to locate an item in a warehouse; it can also not easily store information about the history of its attached object – questions like "where has this parcel been?" are interesting in many applications but require some active participation of the sensor node [246, 392].

Telematics Partially related to logistics applications are applications for the telematics context, where sensors embedded in the streets or roadsides can gather information about traffic conditions at a much finer grained resolution than what is possible today [296]. Such a so-called "intelligent roadside" could also interact with the cars to exchange danger warnings about road conditions or traffic jams ahead.

In addition to these, other application types for WSNs that have been mentioned in the literature include airplane wings and support for smart spaces [245], applications in waste water treatment plants [367], instrumentation of semiconductor processing chambers and wind tunnels [392], in "smart kindergartens" where toys interact with children [783], the detection of floods [88], interactive museums [667], monitoring a bird habitat on a remote island [534], and implanting sensors into the human body (for glucose monitoring or as retina prosthesis) [745]

While most of these applications are, in some form or another, possible even with today's technologies and without wireless sensor networks, all current solutions are "sensor starved" [667]. Most applications would work much better with information at higher spatial and temporal resolution about their object of concern than can be provided with traditional sensor technology. wireless sensor networks are to a large extent about providing the required information at the required accuracy in time with as little resource consumption as possible.

1.3 Types of applications

Many of these applications share some basic characteristics. In most of them, there is a clear difference between **sources** of data – the actual nodes that sense data – and **sinks** – nodes where the data should be delivered to. These sinks sometimes are part of the sensor network itself; sometimes they are clearly systems "outside" the network (e.g. the firefighter's PDA communicating with a WSN). Also, there are usually, but not always, more sources than sinks and the sink is oblivious or not interested in the identity of the sources; the data itself is much more important.

The **interaction patterns** between sources and sinks show some typical patterns. The most relevant ones are:

Event detection Sensor nodes should report to the sink(s) once they have detected the occurrence of a specified event. The simplest events can be detected locally by a single sensor node in isolation (e.g. a temperature threshold is exceeded); more complicated types of events require the collaboration of nearby or even remote sensors to decide whether a (composite) event has occurred (e.g. a temperature gradient becomes too steep). If several different events can occur, **event classification** might be an additional issue.

Periodic measurements Sensors can be tasked with periodically reporting measured values. Often, these reports can be triggered by a detected event; the reporting period is application dependent.

Function approximation and edge detection The way a physical value like temperature changes from one place to another can be regarded as a function of location. A WSN can be used to approximate this unknown function (to extract its spatial characteristics), using a limited number of samples taken at each individual sensor node. This approximate mapping should be made available at the sink. How and when to update this mapping depends on the application's needs, as do the approximation accuracy and the inherent trade-off against energy consumption.

Similarly, a relevant problem can be to find areas or points of the same given value. An example is to find the isothermal points in a forest fire application to detect the border of the actual fire. This can be generalized to finding "edges" in such functions or to sending messages along the boundaries of patterns in both space and/or time [274].

Tracking The source of an event can be mobile (e.g. an intruder in surveillance scenarios). The WSN can be used to report updates on the event source's position to the sink(s), potentially with estimates about speed and direction as well. To do so, typically sensor nodes have to cooperate before updates can be reported to the sink.

These interactions can be scoped both in time and in space (reporting events only within a given time span, only from certain areas, and so on). These requirements can also change dynamically overtime; sinks have to have a means to inform the sensors of their requirements at runtime. Moreover, these interactions can take place only for one specific request of a sink (so-called "one-shot queries"), or they could be long-lasting relationships between many sensors and many sinks.

The examples also have shown a wide diversity in **deployment options**. They range from well-planned, fixed deployment of sensor nodes (e.g. in machinery maintenance applications) to random deployment by dropping a large number of nodes from an aircraft over a forest fire. In addition, sensor nodes can be mobile themselves and compensate for shortcomings in the deployment process by moving, in a postdeployment phase, to positions such that their sensing tasks can be better fulfilled [17]. They could also be mobile because they are attached to other objects (in the logistics applications, for example) and the network has to adapt itself to the location of nodes.

The applications also influence the available **maintenance options**: Is it feasible and practical to perform maintenance on such sensors – perhaps even required in the course of maintenance on associated machinery? Is maintenance irrelevant because these networks are only deployed in a strictly ad hoc, short-term manner with a clear delimitation of maximum mission time (like in disaster recovery operations)? Or do these sensors have to function unattended, for a long time, with no possibility for maintenance?

Closely related to the maintenance options are the **options for energy supply**. In some applications, wired power supply is possible and the question is mute. For self-sustained sensor nodes, depending on the required mission time, energy supply can be trivial (applications with a few days of usage only) or a challenging research problem, especially when no maintenance is possible but nodes have to work for years. Obviously, acceptable price and size per node play a crucial role in designing energy supply.

1.4 Challenges for WSNs

Handling such a wide range of application types will hardly be possible with any single realization of a WSN. Nonetheless, certain common traits appear, especially with respect to the characteristics and the required mechanisms of such systems. Realizing these characteristics with new mechanisms is the major challenge of the vision of wireless sensor networks.

1.4.1 Characteristic requirements

The following characteristics are shared among most of the application examples discussed above:

Type of service The service type rendered by a conventional communication network is evident – it moves bits from one place to another. For a WSN, moving bits is only a means to an end, but not the actual purpose. Rather, a WSN is expected to provide meaningful information and/or actions about a given task: "People want answers, not numbers" (Steven Glaser, UC Berkeley, in [367]). Additionally, concepts like *scoping* of interactions to specific geographic regions or to time intervals will become important. Hence, new paradigms of using such a network are required, along with new interfaces and new ways of thinking about the service of a network.

Quality of Service Closely related to the type of a network's service is the quality of that service. Traditional quality of service requirements – usually coming from multimedia-type applications – like bounded delay or minimum bandwidth are irrelevant when applications are tolerant to latency [26] or the bandwidth of the transmitted data is very small in the first place. In some cases, only occasional delivery of a packet can be more than enough; in other cases, very high reliability requirements exist. In yet other cases, delay *is* important when actuators are to be controlled in a real-time fashion by the sensor network. The packet delivery ratio is an insufficient metric; what is relevant is the amount and quality of information that can be extracted at given sinks about the observed objects or area.

Therefore, adapted quality concepts like reliable detection of events or the approximation quality of a, say, temperature map is important.

Fault tolerance Since nodes may run out of energy or might be damaged, or since the wireless communication between two nodes can be permanently interrupted, it is important that the WSN as a whole is able to tolerate such faults. To tolerate node failure, redundant deployment is necessary, using more nodes than would be strictly necessary if all nodes functioned correctly.

Lifetime In many scenarios, nodes will have to rely on a limited supply of energy (using batteries). Replacing these energy sources in the field is usually not practicable, and simultaneously, a WSN must operate at least for a given mission time or as long as possible. Hence, the **lifetime** of a WSN becomes a very important figure of merit. Evidently, an energy-efficient way of operation of the WSN is necessary.

As an alternative or supplement to energy supplies, a limited power source (via power sources like solar cells, for example) might also be available on a sensor node. Typically, these sources are not powerful enough to ensure continuous operation but can provide some recharging of batteries. Under such conditions, the lifetime of the network should ideally be infinite.

The lifetime of a network also has direct trade-offs against quality of service: investing more energy can increase quality but decrease lifetime. Concepts to harmonize these trade-offs are required.

The precise *definition of lifetime* depends on the application at hand. A simple option is to use the time until the first node fails (or runs out of energy) as the network lifetime. Other options include the time until the network is disconnected in two or more partitions, the time until 50 % (or some other fixed ratio) of nodes have failed, or the time when for the first time a point in the observed region is no longer covered by at least a single sensor node (when using redundant deployment, it is possible and beneficial to have each point in space covered by several sensor nodes initially).

Scalability Since a WSN might include a large number of nodes, the employed architectures and protocols must be able scale to these numbers.

Wide range of densities In a WSN, the number of nodes per unit area – the *density* of the network – can vary considerably. Different applications will have very different node densities. Even within a given application, density can vary over time and space because nodes fail or move; the density also does not have to homogeneous in the entire network (because of imperfect deployment, for example) and the network should adapt to such variations.

Programmability Not only will it be necessary for the nodes to process information, but also they will have to react flexibly on changes in their tasks. These nodes should be programmable, and their programming must be changeable during operation when new tasks become important. A fixed way of information processing is insufficient.

Maintainability As both the environment of a WSN and the WSN itself change (depleted batteries, failing nodes, new tasks), the system has to adapt. It has to monitor its own health and status to change operational parameters or to choose different trade-offs (e.g. to provide lower quality when energy resource become scarce). In this sense, the network has to maintain itself; it could also be able to interact with external maintenance mechanisms to ensure its extended operation at a required quality [534].

1.4.2 Required mechanisms

To realize these requirements, innovative mechanisms for a communication network have to be found, as well as new architectures, and protocol concepts. A particular challenge here is the need to find mechanisms that are sufficiently specific to the idiosyncrasies of a given application to support the specific quality of service, lifetime, and maintainability requirements [246]. On the other hand, these mechanisms also have to generalize to a wider range of applications lest a complete from-scratch development and implementation of a WSN becomes necessary for every individual application – this would likely render WSNs as a technological concept economically infeasible.

Some of the mechanisms that will form typical parts of WSNs are:

Multihop wireless communication While wireless communication will be a core technique, a direct communication between a sender and a receiver is faced with limitations. In particular, communication over long distances is only possible using prohibitively high transmission power. The use of intermediate nodes as relays can reduce the total required power. Hence, for many forms of WSNs, so-called *multihop communication* will be a necessary ingredient.

Energy-efficient operation To support long lifetimes, energy-efficient operation is a key technique. Options to look into include energy-efficient data transport between two nodes (measured in J/bit) or, more importantly, the energy-efficient determination of a requested information. Also, nonhomogeneous energy consumption – the forming of “hotspots” – is an issue.

Auto-configuration A WSN will have to configure most of its operational parameters autonomously, independent of external configuration – the sheer number of nodes and simplified deployment will require that capability in most applications. As an example, nodes should be able to determine their geographical positions only using other nodes of the network – so-called “self-location”. Also, the network should be able to tolerate failing nodes (because of a depleted battery, for example) or to integrate new nodes (because of incremental deployment after failure, for example).

Collaboration and in-network processing In some applications, a single sensor is not able to decide whether an event has happened but several sensors have to collaborate to detect an event and only the joint data of many sensors provides enough information. Information is processed in the network itself in various forms to achieve this collaboration, as opposed to having every node transmit all data to an external network and process it “at the edge” of the network.

An example is to determine the highest or the average temperature within an area and to report that value to a sink. To solve such tasks efficiently, readings from individual sensors can be *aggregated* as they propagate through the network, reducing the amount of data to be transmitted and hence improving the energy efficiency. How to perform such aggregation is an open question.

Data centric Traditional communication networks are typically centered around the transfer of data between two specific devices, each equipped with (at least) one network address – the operation of such networks is thus **address-centric**. In a WSN, where nodes are typically deployed redundantly to protect against node failures or to compensate for the low quality of a single node's actual sensing equipment, the identity of the particular node supplying data becomes irrelevant. What is important are the answers and values themselves, not which node has provided them. Hence, switching from an address-centric paradigm to a **data-centric** paradigm in designing architecture and communication protocols is promising.

An example for such a data-centric interaction would be to request the average temperature in a given location area, as opposed to requiring temperature readings from individual nodes. Such a data-centric paradigm can also be used to set conditions for alerts or events ("raise an alarm if temperature exceeds a threshold"). In this sense, the data-centric approach is closely related to query concepts known from databases; it also combines well with collaboration, in-network processing, and aggregation.

Locality Rather a design guideline than a proper mechanism, the principle of locality will have to be embraced extensively to ensure, in particular, scalability. Nodes, which are very limited in resources like memory, should attempt to limit the state that they accumulate during protocol processing to only information about their direct neighbors. The hope is that this will allow the network to scale to large numbers of nodes without having to rely on powerful processing at each single node. How to combine the locality principle with efficient protocol designs is still an open research topic, however.

Exploit trade-offs Similar to the locality principle, WSNs will have to rely to a large degree on exploiting various inherent trade-offs between mutually contradictory goals, both during system/protocol design and at runtime. Examples for such trade-offs have been mentioned already: higher energy expenditure allows higher result accuracy, or a longer lifetime of the entire network trades off against lifetime of individual nodes. Another important trade-off is node density: depending on application, deployment, and node failures at runtime, the density of the network can change considerably – the protocols will have to handle very different situations, possibly present at different places of a single network. Again, not all the research questions are solved here.

Harnessing these mechanisms such that they are easy to use, yet sufficiently general, for an application programmer is a major challenge. Departing from an address-centric view of the network requires new programming interfaces that go beyond the simple semantics of the conventional socket interface and allow concepts like required accuracy, energy/accuracy trade-offs, or scoping.

1.5 Why are sensor networks different?

On the basis of these application examples and main challenges, two close relatives of WSNs become apparent: Mobile Ad Hoc Networks (MANETs) on the one hand and fieldbuses on the other hand.

1.5.1 Mobile ad hoc networks and wireless sensor networks

An ad hoc network is a network that is setup, literally, for a specific purpose, to meet a quickly appearing communication need. The simplest example of an ad hoc network is perhaps a set of computers connected together via cables to form a small network, like a few laptops in a meeting room. In this example, the aspect of *self-configuration* is crucial – the network is expected to work without manual management or configuration.

Usually, however, the notion of a MANET is associated with wireless communication and specifically *wireless* multihop communication; also, the name indicates the mobility of participating nodes as a typical ingredient. Examples for such networks are disaster relief operations – firefighters communicate with each other – or networks in difficult locations like large construction sites, where the deployment of wireless infrastructure (access points etc.), let alone cables, is not a feasible option. In such networks, the individual nodes together form a network that relays packets between nodes to extend the reach of a single node, allowing the network to span larger geographical areas than would be possible with direct sender – receiver communication. The two basic challenges in a MANET are the reorganization of the network as nodes move about and handling the problems of the limited reach of wireless communication. Literature on MANETs that summarize these problems and their solutions abound, as these networks are still a very active field of research; popular books include [635, 793, 827].

These general problems are shared between MANETs and WSNs. Nonetheless, there are some principal differences between the two concepts, warranting a distinction between them and regarding separate research efforts for each one.

Applications and equipment MANETs are associated with somewhat different applications as well as different user equipment than WSNs: in a MANET, the terminal can be fairly powerful (a laptop or a PDA) with a comparably large battery. This equipment is needed because in the typical MANET applications, there is usually a human in the loop: the MANET is used for voice communication between two distant peers, or it is used for access to a remote infrastructure like a Web server. Therefore, the equipment has to be powerful enough to support these applications.

Application specific Owing to the large number of conceivable combinations of sensing, computing, and communication technology, many different application scenarios for WSNs become possible. It is unlikely that there will be a “one-size-fits-all” solution for all these potentially very different possibilities. As one example, WSNs are conceivable with very different network densities, from very sparse to very dense deployments, which will require different or at least adaptive protocols. This diversity, although present, is not quite as large in MANETs.

Environment interaction Since WSNs have to interact with the environment, their traffic characteristics can be expected to be very different from other, human-driven forms of networks. A typical consequence is that WSNs are likely to exhibit very low data rates over a large timescale, but can have very bursty traffic when something happens (a phenomenon known from real-time systems as event showers or alarm storms). Long periods (months) of inactivity can alternate with short periods (seconds or minutes) of very high activity in the network, pushing its capacity to the limits. MANETs, on the other hand, are used to support more conventional applications (Web, voice, and so on) with their comparably well understood traffic characteristics.

Scale Potentially, WSNs have to scale to much larger numbers (thousands or perhaps hundreds of thousands) of entities than current ad hoc networks, requiring different, more scalable solutions. As a concrete case in point, endowing sensor nodes with a unique identifier is costly (either at production or at runtime) and might be an overhead that could be avoided – hence, protocols that work without such identifiers might become important in WSNs, whereas it is fair to assume such identifiers to exist in MANET nodes.

Energy In both WSNs and MANETs, energy is a scare resource. But WSNs have tighter requirements on network lifetime, and recharging or replacing WSN node batteries is much less an option than in MANETs. Owing to this, the impact of energy considerations on the entire system architecture is much deeper in WSNs than in MANETs.

Self configurability Similar to ad hoc networks, WSNs will most likely be required to self-configure into connected networks, but the difference in traffic, energy trade-offs, and so forth, could require new solutions. Nevertheless, it is in this respect that MANETs and WSNs are probably most similar.

Dependability and QoS The requirements regarding dependability and QoS are quite different. In a MANET, each individual node should be fairly reliable; in a WSN, an individual node is next to irrelevant. The quality of service issues in a MANET are dictated by traditional applications (low jitter for voice applications, for example); for WSNs, entirely new QoS concepts are required, which also take energy explicitly into account.

Data centric Redundant deployment will make data-centric protocols attractive in WSNs. This concept is alien to MANETs. Unless applications like file sharing are used in MANETs, which do bear some resemblance to data centric approaches, data-centric protocols are irrelevant to MANETs – but these applications do not represent the typically envisioned use case.

Simplicity and resource scarceness Since sensor nodes are simple and energy supply is scarce, the operating and networking software must be kept orders of magnitude simpler compared to today's desktop computers. This simplicity may also require breaking with conventional layering rules for networking software, since layering abstractions typically cost time and space. Also, resources like memory, which is relevant for comparably heavy-weight routing protocols as those used in MANETs, is not available in arbitrary quantities, requiring new, scalable, resource-efficient solutions.

Mobility The mobility problem in MANETs is caused by nodes moving around, changing multihop routes in the network that have to be handled. In a WSN, this problem can also exist if the sensor nodes are mobile in the given application. There are two additional aspects of mobility to be considered in WSNs.

First, the sensor network can be used to detect and observe a physical phenomenon (in the intrusion detection applications, for example). This phenomenon is the cause of events that happen in the network (like raising of alarms) and can also cause some local processing, for example, determining whether there really is an intruder. What happens if this phenomenon moves about? Ideally, data that has been gathered at one place should be available at the next one. Also, in tracking applications, it is the explicit task of the network to ensure that some form of activity happens in nodes that surround the phenomenon under observation.

Second, the sinks of information in the network (nodes where information should be delivered to) can be mobile as well. In principle, this is no different than node mobility in the general MANET sense, but can cause some difficulties for protocols that operate efficiently in fully static scenarios. Here, carefully observing trade-offs is necessary.

Furthermore, in both MANET and WSNs, mobility can be correlated – a group of nodes moving in a related, similar fashion. This correlation can be caused in a MANET by, for example, belonging to a group of people traveling together. In a WSN, the movement of nodes can be correlated because nodes are jointly carried by a storm, a river, or some other fluid.

In summary, there are commonalities, but the fact that WSNs have to support very different applications, that they have to interact with the physical environment, and that they have to carefully adjudicate various trade-offs justifies considering WSNs as a system concept distinct from MANETs.

1.5.2 Fieldbuses and wireless sensor networks

Fieldbuses are networks that are specifically designed for operation under hard real-time constraints and usually with inbuilt fault tolerance, to be used predominantly in control applications, that is, as part of a control loop. Examples include the Profibus and IEEE 802.4 Token Bus networks [372] for factory floor automation or the CAN bus for onboard networks in cars; some example summaries on the topic include [532, 644, 881]. Because of the stringent hard real-time requirements, these networks are usually wired and only the layers one (physical), two (link layer), and seven (application) of the OSI reference model are used, avoiding communication over multiple hops and associated queuing delays in intermediate nodes. Nevertheless, a number of research efforts deal with realizing fieldbus semantics on top of wireless communication, despite its inherently limited error rates that jeopardize real-time guarantees [200, 687, 878].

Since fieldbuses also have to deal with the physical environment for which they report sensing data and which they control, they are in this sense very similar to WSNs. With some justification, WSNs can be considered examples of wireless fieldbuses. Some differences do exist, however: WSNs do mostly not attempt to provide real-time guarantees in the range of (tens of) milliseconds but are rather focused on applications that can tolerate longer delays and some jitter (delay variability). Also, the adaptive trade-offs that WSNs are willing to make (accuracy against energy efficiency, for example) is a concept that is not commonly present in the fieldbus literature; specifically, fieldbuses make no attempt to conserve energy, and their protocols are not prepared to do so.

But these distinctions can only serve as a rough guideline; the borderline between these two research areas is certainly a blurry one.

1.6 Enabling technologies for wireless sensor networks

Building such wireless sensor networks has only become possible with some fundamental advances in enabling technologies. First and foremost among these technologies is the miniaturization of hardware. Smaller feature sizes in chips have driven down the power consumption of the basic components of a sensor node to a level that the constructions of WSNs can be contemplated. This is particularly relevant to microcontrollers and memory chips as such, but also, the radio modems, responsible for wireless communication, have become much more energy efficient. Reduced chip size and improved energy efficiency is accompanied by reduced cost, which is necessary to make redundant deployment of nodes affordable.

Next to processing and communication, the actual sensing equipment is the third relevant technology. Here, however, it is difficult to generalize because of the vast range of possible sensors – Chapter 2 will go more into details here.

These three basic parts of a sensor node have to accompanied by power supply. This requires, depending on application, high capacity batteries that last for long times, that is, have only a negligible self-discharge rate, and that can efficiently provide small amounts of current. Ideally, a sensor node also has a device for **energy scavenging**, recharging the battery with energy gathered from the environment – solar cells or vibration-based power generation are conceivable options. Such a concept requires the battery to be efficiently chargeable with small amounts of current, which is not a standard ability. Both batteries and energy scavenging are still objects of ongoing research.

The counterpart to the basic hardware technologies is software. The first question to answer here is the principal division of tasks and functionalities in a single node – the architecture of the operating system or runtime environment. This environment has to support simple retasking, cross-layer information exchange, and modularity to allow for simple maintenance. This software architecture on a single node has to be extended to a network architecture, where the division of tasks between nodes, not only on a single node, becomes the relevant question – for example, how to structure interfaces for application programmers. The third part to solve then is the question of how to design appropriate communication protocols.

This book only touches briefly on the hardware aspects of WSNs. It is also not much concerned with the questions of appropriate runtime environments. It focuses, rather, on the WSNs architecture and protocols to solve the communication questions as such.