When HART Goes Wireless: Understanding and Implementing the WirelessHART Standard

Anna N. Kim*‡ Fredrik Hekland† Stig Petersen§ Paula Doyle‡
Centre for Quantifiable Quality of Service in Communication Systems (Q2S)
Norwegian University of Science and Technology, Trondheim, Norway
†ABB Corporate Research Centre (CRC),§ Sintef ICT Trondheim, Norway
‡ ABB Strategic R&D for Oil and Gas, Oslo Norway

Abstract

As a newly released industrial communication standard, WirelessHART complements the ever so successful HART field devices by providing the possible means for communicating via wireless channels. The WirelessHART standard is designed to offer simple configuration, flexible installation and easy access of instrument data, and at the same time, ensure robust and reliable communications. In this paper, we first look closely into the specifications and present a comprehensive overview of the standard by summarizing the main functions of the various protocol layers. We then survey the literature and identify amongst the existing methods and algorithms, which ones can be effectively adopted in implementing the standard. More specifically, we set our focus on issues relating to realization of the medium access layer and the network manager, which are essential in creating a successful WirelessHART network for specific applications.

1. Introduction

The adoption of wireless technology has been slow in the process automation and manufacturing industries. A major reason for this has been the lack of an open standard which both fulfills the industrial requirements as well as ensures that the customers are not locked to one single supplier. This is about to change, at least for the process automation industries. Owing to the HART Communication Foundation and its member companies' efforts to create a wireless interface for the HART standard, the benefits of wireless access to field instruments might finally outweigh the risks and uncertainties of rolling out a wireless network to the field devices.

The HART field communication protocol has more than 20 years on its back and is still going strong. With an estimated 24 million devices in operation, the standard has proven its worth and shows no sign reduction is sales either. Although part of the explanation of HART's popularity is the failure to create a single standard for a digital fieldbus, the major motivation for customers to still choose HART in 2008 is its inherent simplicity and robustness. The combination of the analog 4-20mA control loop with a superimposed digital signal for configuration and diagnostics simply works, and requires much less training than the competing digital fieldbuses. Another drawback of the digital fieldbuses is the higher power consumption due to higher complexity.

With the release of Version 7 of the HART protocol in September 2007, the process automation industry now has access to an open standard which offered a wireless interface to field devices, referred to as WirelessHART. WirelessHART promises to bring the heritage of simplicity and robustness the customers know from the earlier revision of the HART standard. Little or no training is necessary for the plant workers to start using the wireless products as the wireless mesh network is self-organizing. Robust communication is expected through the application of modulation techniques like both direct-sequence spread-spectrum (DSSS) and frequency-hopping spread-spectrum (FHSS), as well as by retranmission mechanisms and spatial path diversity through the mesh network. Proper data security is also covered in the standard; a multi-layered approach for authentication, integrity and encryption using well-tested encryption algorithms ensures the user can select the level of security necessary for the plant.

Communication standards that are similar to WirelessHART include Zigbee [3] and ISA 100 [4]. The ZigBee specification is a low rate, low power wireless mesh networking standard developed by the ZigBee Alliance, primarily targeting home automation and consumer electronics applications. Initially released in 2004 and updated in 2006, the specification provides a network and application layer on top of the PHY and MAC layers of the IEEE 802.15.4 specification.

*Centre for Quantifiable QoS in Communication Systems, Centre of Excellence is appointed by the Research Council of Norway (NFR) and funded by NFR, NTNU and Uninett.

The author(s) wish to acknowledge the support of NFR and the TAIL IO project for their continued funding and support for this research. The TAIL IO project is an international cooperative research project led by StatoilHydro and an R&D consortium consisting of ABB, IBM, Aker Kvaerner and SKF

1-4244-1506-3/08/$25.00 ©2008 IEEE

Unlike WirelessHART, ZigBee does not support frequency hopping. A ZigBee network operates on the same static channel throughout its entire lifetime. The use of a static channel makes the ZigBee networks more susceptible to noise and interference, and because of this, ZigBee has not been regarded as robust enough for harsh radio frequency environments often encountered in industrial applications. To counter this, and to gain more momentum as a viable option for industrial instrumentation, the ZigBee Alliance released the ZigBee PRO specification in October 2007. The ZigBee PRO is specifically aimed at the industrial market, employing, among other things, enhanced security features as well as a new "frequency agility" concept which allows for an entire network to change its operative channel when faced with reduced link qualities caused by noise and/or interference. Frequency hopping, which is a more flexible solution that the frequency agility, requires modifications to the IEEE 802.15.4 MAC layer, and as the ZigBee Alliance want to fully adopt the IEEE 802.15.4 specification, this is not a viable option for ZigBee PRO.

ISA100 on the other hand, is more of a closer contestant for industrial communications. The goal of the ISA100 standards committee is to create a family of wireless standards for industrial automation. The first standard to emerge will be the ISA100.11a Release 1, which is expected to be ratified by the fourth quarter of 2008. The aim of the ISA100.11a is to provide secure and reliable wireless communication for fixed, portable and moving devices for non-critical monitoring and control applications.

The main difference between WirelessHART and ISA100 lies on the application layer, as ISA100 is designed for handling, in addition to HART commands, also Fieldbus Foundation, Profibus, and Modbus. Furthermore, the ISA100.11 a incorporates management functions which support management in five areas across the network and across all layers of the architecture. The five management areas are accounting, configuration, fault, performance and security. The management service includes a device management application process that resides on all ISA100.11a devices, as well as one or more system manager applications that reside on a small subset of devices.

Although the use of the WirelessHART standard is supposed to be easy, the implementation is more complex than the wired counterpart. Proper design of all aspects of the system is needed to ensure long network lifetime with good stability. This requires more effort from the device manufacturers, but also enables manufacturers to distinguish themselves through offering better implementations than their competitors.

In this paper we first walk through the central building blocks of the standard, and look at the medium-access and network layers which are important from a system perspective. We then survey the literature and identify some existing algorithms and methods that can be used for solving typical problems that are associated with implementing the standard. We conclude the paper with possible improvements and extensions for the standard and give suggestions on future directions.

2 The WirelessHART Communication Protocol

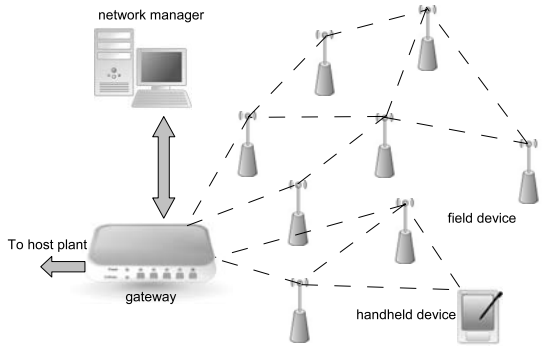

Figure 1. A generic WirelessHART network

Figure 1 depicts a generic WirelessHART network. It is formed by a group of *network devices*. They can either be a field device which is directly connected to the process plant, or handheld devices from for example a maintenance worker. As shown in the figure, the network supports both direct connection between device and gateway (star topology) and connections over multiple hops (mesh topology). Each network device therefore must be able to function as source, sink and router.

The WirelessHART gateway acts as the bridge connecting the WirelessHART network to the process plant. The basic schematics are illustrated in Figure 2. It consists of a *virtual gateway* and one or more network access points. Host applications access the network devices through the service interface, which can have single or multiple ports. Operators can also monitor or configure particular field device through the process plant backbone. The network access points provide the actual physical connection to the WirelessHART network. The virtual gateway works as the sink and source for the network traffic. It is required to be a HART type device, namely one that supports all HART application commands and also able to translate cached data that can be interpreted by the host applications. The gateway also provides buffering for bursty and large data transfers, command responses, event notification and diagnostics.

The virtual gateway communicates directly with the network manager, which is responsible for the configuration and maintenance of the WirelessHART network. Each network consists of one and *only one* network manager. The manager requests information from field devices via the gateway to decide for example, how the communication routes should be setup. The host application can also provide input to the network manager when for ex-

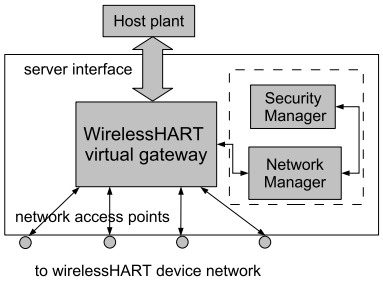

Figure 2. Schematics of a possible single box WirelessHART gateway implementation with multiple network access points. Note the network manager and security manager need not to be in the same box physically.

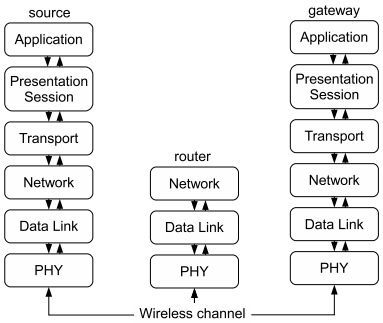

Figure 3. Communication between source device and the gateway through a routing device, modeled by OSI layers

ample, data from a specific device is needed and prioritized scheduling is preferred. Details regarding the network manager will be described in the following section.

The security manager works together with the network manager to prevent possible intrusion and attacks to the WirelessHART network. Multiple networks can be connected to one security manager. It generates session keys, joint keys and network key, which are further propagated to the field devices by the network manager.

The WirelessHART standard specifies the communication protocol stack using the OSI model, as depicted in Figure 3. At the bottom, there is the physical layer which is responsible for signaling, modulation and actual transmission of data. Above that, there is the data link layer which determines how the common wireless medium is shared between the network devices. It is also responsible for formatting the data packets, detection/correction of bit errors. The network layer is the core of the WirelessHART standard. Its responsibilities include routing, topology control, end-to-end security and session management. The transport layer ensures the end-to-end transmission reliability and flow control. Just like HART, the WirelessHART have command oriented application layer. In addition to define data type and formats, all commands to the lower layers and the gateway are generated here.

When for example, a field device's task is to collect and transmit temperature measurement to the gateway, the measurement data is collected and pass "down" through the OSI layer protocol stack and communicated over the wireless channel. If transmission involves a intermediate device acts as a router, the packets only go up to (and down from) the network layer. At its destination, data packets travel "up" the protocol stack and finally handled by the application layer at the gateway. Successful transmission therefore relies on proper functioning of each layer. On the network level, stability and functionality are ensured by coordination of devices and intelligent allocation of network resources. These tasks are carried out jointly by the responsible layers over all devices.

Since most details regarding the actual standard can be obtained through the HART communication foundation [1], we will only summarize how the each layer functions and highlight the areas which is either unique to WirelessHART or essential in delivering the services it is designed for.

2.1 PHY and MAC Layer

The wirelssHART PHY layer is based on the IEEE STD 802.15.4-2006 [2]. It employs O-QPSK (Offset Quadrature Phase Shift Keying) and operates on the 2.4GHz unlicensed band with data rate up to 250 kbps. DSSS is utilized to resist interference from jamming. It is combined with FHSS, where the radio carrier hops over multiple frequency bands(channels) using a pseudo random sequence. FHSS is effective in overcome narrow band interference such as that from multipath fading.

At the MAC layer, which is a sublayer of data link layer, WirelessHART utilizes time division multiple access (TDMA) to ensure contention free transmission. Each time-slot is of 10 msec duration. Between two communicating devices, the time slot has just sufficient room for transmit/receive one data unit plus an acknowledgement packet. In case of broadcast message, acknowledgement is no longer necessary and multiple receivers can be assigned to the same slot. A collection of time slots form a *superframe*. The size of superframe may vary and all the devices in the network are required to support multiple concurrent frames. The superframes are then repeated at a fixed rate throughout the network lifetime.

In essence, communication in a WirelessHART network is defined through the superframe, time slot and the wireless links. The link (aside from broadcast links) pairs two device. It can belong to *one and only one* superframe. Each link is associated with a set of channels used for frequency hopping. All devices in the network share an identical channel lists indicating which channel can be used. Each network device also maintains a list of links, showing their connection to other devices. Upon transmission, the device randomly chooses a link from the available link list for transmission and uses the channel-offset to calculate the link frequency/channel. Transmission is then initiated at the designated time slot within a superframe. This

process repeats at the next designated time slot, which may be in the same superframe, or a concurrent one. Due to channel diversity provided by frequency hopping, the time slot can also be shared by different nodes, which is used for for example advertising device presence or neighbour discovery. In case of collision at the destination node, the source devices use the random back-off mechanism to wait for new opportunities of transmission. For each unacknowledge transmission, the waiting window size is increased. This is similar to other contention based channel access schemes. Broadcast messages however, are not allowed on shared time slots.

As each device supports a class of data units, these units also different priority. In general, commands, which include control, configuration information and network related diagnostics are classified to have highest priority. They are followed by process data which contain process information (measurements) or network statistics. The lowest priority packets belong to the "alarm" class which contains only alarm and event payload. All other packets are then grouped into the "normal" class which has priority over "alarm". Priority classification is useful for flow control, which decides how many of which type of packets can be buffered at a relay device.

2.2 Network Manager

The network manager maintains the health and well-being of a WirelessHART network and is responsible for a range of functions which are normally categorized separately in the data link layer and network layer. We describe the network manager as a single unit here.

When the WirelessHART network is first initialized, a unique network ID along with security keys from the security manager are provided to the network manager. It establishes the connection with the gateway and the network access points to secure bandwidth needed for management and control packets going to and from the network devices. When a new device wishes to join the network, the manager validates its integrity using join keys and the network ID to ensure it is trustworthy and is joining the right network. Once authenticated, the network manager provides the device with necessary network and session keys from the security manager and assigns a 16bit network address.

The network manager maintains a complete list of all devices and has full knowledge of the network topology. It also responds to host application requests regarding any network level information.

The transmission routes from source to destination of the entire network are configured by the network manager. Each network device maintains its own neighbour table, which contains a list of all devices it is able to connect to. The network manager can, for instance, pull the neigbour tables from each devices within the network, along with for example the current battery status to construct optimal routes. The result is a collection of *routing graphs* where each edge of the graph represents a possible transmission link between two devices. The routing graphs are not unique and can therefore overlap. They are however unidirectional and each graph is associated with a unique graph ID. This ID information is then passed down to the devices to be placed in the packet network header to determine which route to be used for transmission. It is required that the device has at least two neighbours for transmission in each routing graph to ensure path diversity hence better reliability.

Alternatively, the network manager can also configure a list of devices which are capable of *source routing*. The sending device (source) specifies a single path route to the receiving device (destination) using a list of device addresses. The address field is limited to 4 and is used primarily for testing and trouble shooting of network path.

Another important network manager responsibility is link scheduling. While graph routing determines *where* the packets should be sent, link scheduling tells *when* the packet should be sent. Since WirelessHART utilizes TDMA, each transmission link is associated with a defined time slot. Properly configured the link schedules can reduce the latency, increase network throughput and balance the network load. The network manager creates and maintains a network wide link table. It specifies the time slot and superframe the link associated with, along with its type, function (transmit, receive or shared), neighbour information, channel offset and the devices it connects. The network manager can activate and deactivate superframes when necessary and keeps track of blacklisted channels.

The network manager is also responsible for collecting system performance and diagnostic information. Such information can be used to monitor and assess the overall state of the network. Real time evaluation can be made possible through the ASN (absolute slot number) Snippet field, to identify for example the age of a packet. When the network environment changes, for example severe and abrupt interference, a number of failed or newly joint devices, the network manager must adapt the network operation by updating the routing and scheduling information accordingly.

2.3 Transport Layer

A unique feature of the WirelessHART transport layer is the *block data transfer mechanism*. It sets up a connection oriented communication link between the host application and the field device. The host application can configure the slave device by opening a port onboard the device using a HART command. The port specifications are also part of the WirelessHART standard. Once the port is opened, transmission rate between the device and host application is negotiated with the network manager to maximize throughput. The block data transfer is required to be reliable and end-to-end acknowledgement is necessary to keep track of the data stream. This may call for the network manager to update its routing and scheduling plan to provide the necessary priority.

On a more general basis, the transport layer of Wire-

lessHART is analogous to that of the Internet. It supports both TCP/IP like transfer with acknowledgement which is useful for event notification and UDP like without acknowledgements when sending real time process data that has shorter life span. With acknowledgement service, automatic repeat request (ARQ) can be utilized to ensure end-to-end delivery of messages. The default number of re-transmission is set to 5.

2.4 Summary

The WirelessHART standard can be summarized in regards to the following important aspects.

- The WirelessHART network has a *centralized* operation architecture. The network manager configures and maintains the network by gathering information from all devices and giving operation instructions. The field device is, in comparison, much less intelligent. It has only knowledge of local information and is only responsible for link level transmission/reception.

- Energy efficiency of a WirelessHART network relies on, to a great extend, the design of the network manager. In a mesh topology, routing and link scheduling will directly affect how quickly the power of a device is drained when it is often used for forwarding data. It is important to avoid situations where some devices become isolated and cut-off from the rest of the network. Related issue, such as load balancing, will be addressed in the following section.

- Reliable data transmission is ensured through a combination of PHY layer, MAC, network and transport layer mechanisms. That is, FHSS, DSSS to combat the fading and interference in the wireless channel, link level ACK to trigger retransmission when needed, path diversity in graphic routing and transport protocol with end-to-end acknowledgements.

- Factors influence the latency of data transmission include routing/topology, link scheduling and the network size. Since WirelessHART support both star and mesh topology, direct connection to the gateway can reduce transmission time although may come at the expense of higher power consumption. Under the mesh topology, on the other hand, latency can still be minimized with smart routing, form example limiting the maximum number of hops and intelligent scheduling where higher priority can be set to certain links to ensure quick and continuous data flow.

In the following section, we will review a class of methods and algorithms that can be utilized to realize the protocols within the standard specifications, while at the same time taking the constraints from above mentioned aspects into consideration.

3 Implementing WirelessHART

Effectively implementing the WirelessHART standard is not a simple task. On one hand, the standard states clearly the basic requirement for the various protocols and devices. At the same time it must also leave sufficient room for different realization schemes. Aside from supporting/allowing WirelessHART compliant products from various vendors. The intended diverse array of applications such as process monitoring, asset management, process control, health safety and environmental monitoring, becomes another motivation for such flexibility.

In this section, we address a few important issues in implementing WirelessHART. In particular, we focus on the MAC and the network layers, as they contribute most to the functionalities of WirelessHART network and are loosely defined. The PHY layer on the other hand is more restricted. There are already a group of off-the-shelf products (e.g. DUST™ Networks and SensiNet®). The reviewed algorithms can be used as guidelines and recommendations to how to optimally design and configure the WirelessHART network according to the application(s).

3.1 MAC Layer Synchronization

Since WirelessHART uses TDMA, network time keeping is essential. Each device must operate on a common time within certain tolerance to ensure proper access of the wireless channels. Multihop communication in a mesh topology also adds another degree of complexity to time synchronization.

Generally speaking, due to imperfections of device hardware, there are two common factors affecting accuracy of local time. One is clock drift, which indicates the frequency of local clock's change over time; and the other is clock offset, which is the difference from real time. The WirelessHART standard outlines a few basic mechanisms for time synchronization. For example, the relative offset at the receiver can be calculated using the time-stamp of received packet relative to its ideal time. This information is then communicated to the sender device via ACK. To combat clock drifts, the source device is expected to transmit "Keep-Alive" signals no more frequent than every 30 seconds when temperature is varying 2° C per minute or less, to its neighbours.

Since the basic time slot in WirelessHART is only 10 msec, accuracy of synchronization is therefore crucial. At the same time, it may not be desirable to overcome clock drift by frequent "Keep-Alive" messages that consume power and flood the network. Other important factors to consider are scalability, robustness, network lifetime, cost and immediacy in emergency and alarm situations [5]. There are a number of synchronization methods that designed for sensor networks which take these aspects into account. For example, RBS (reference broadcast synchronization) introduced in [6] uses a third-party for synchronization instead of synchronizing the sender with a receiver (as suggested in the WirelessHART standard).

Here the device broadcasts a reference beacon (pulse) to its neighbours, the receivers exchange the receiving time information of the pulse to estimate the relative offest. As a result there can be more than two receivers synchronized at the same time and higher precisions can be achieved using more reference pulses. Implementation on Berkeley motes showed precision of $6.29 \pm 6.45 \mu$ seconds. A competing scheme was proposed in [7]. Network synchronization is achieved through a two phase process. First a level discovery phase is used to define a hierarchical topology through broadcast of level discovery packet, with the root node (in our case, the gateway) assigned level 0. Assuming clock drift between two devices A and B (of two different levels) is constant in duration of message exchange, and constant propagation delay, device A sends *synchronization pulse* packet to B, which B responses at a later time. Local time on each device is then used to calculate the relative clock drift and offset. The process is repeated through out the network tree and synchronize all nodes. Experimental results of TPSN on Berkeley motes offered 16.9 μ seconds error and showed that sender-receiver based synchronization is more effective than receiver based (e.g. RBS). Tiny-sync and its extension Mini-sync [8] improve synchronization accuracy by using a triplet of time stamps from two-way messaging to constrain the relative clock drift and offset. Accuracy of estimation ultimately come at the expanse of complexity. With minimizing complexity in mind and uses accuracy instead as a constraint, [9] proposed a light tree-based synchronization. The device decides when it needs to synchronize, with desired accuracy. It then sends a synchronization request to the closest reference node, which then further propagates the request to all device along the route to the gateway until an already synchronized device is reached. This reactive approach provides savings in terms of eliminating unnecessary synchronization initiatives, however may suffer from repeated synchronization on overlapping routes.

3.2 Network Layer Protocols

As we summarized earlier, the network layer protocols of WirelessHART influence network performance factors such as energy efficiency, reliability and latency. Here we set our focus on the two main responsibilities of the network layer, namely routing and link scheduling and highlight the main results.

3.2.1 Routing and topology control

There has been a wide class of routing algorithms proposed for wireless sensor networks [10]. Many are however not suitable in the WirelessHART context, for example the data centric type routing [11] which routes are discovered through a random walk until the event (device is reached) when it is request through data pulling or query. The scale of the WirelessHART network often is not on the order of thousands or more and routing is exclusively determined by the network manager which has overview of the network topology.

More suited for a WirelessHART network, routes can be determined for example using shortest path as an optimization metric. Transmission energy can be modeled as a function of distance between sender and transmitter. Using maximum transmission distance as a constraint, routes can be formed using standard shortest-path (minimum hop) routing algorithm [12]. Simulation results show a tight constraint (short distance) results in a sparse routing graph, which may compromise reliability from path redundancy. Choosing a larger maximum distance constraint leads to a complete graph, however also leads to more direct route to the gateway, which depletes more of transmission power.

Alternatively, one may also address energy efficiency in the sense of maintaining network *survivability/lifetime*, that is, to ensure the network is operational for as long as possible. The rate of energy depletion of power limited devices then should be approximately on the same order. Energy aware routing algorithm in [13] was proposed for this purpose. Instead of always choosing the path that consumes the lowest energy which can lead to quick depletion and broken routes, suboptimal routes are chosen occasionally. This can be achieved by having multiple routes between source and destination, which comes in line with the graph routing requirement in WirelessHART. The energy metric used is based on the residual energy and the energy required for transmit and receive on the link. Again, these are information can be made easily available to the network manager. The protocol can be modified to fit the WirelessHART standard. During the startup phase, instead of having the device calculating the cost of routing using the energy metric, it is up to the network manager to construct the routing table using the cost function. At the data communication phase, again it is the network manager that decides the alternating use of available routes for each device, based on the calculated probability of which results from the cost of the path. Using a set of static nodes distributed over area of 1000m^2, simulation results showed an increased network lifetime of 44% with the proposed routing protocol.

Constructed routes in a network also require maintenance. Consider again energy efficiency as the main objective, the network manager needs to continuously monitor the available energy level of the field devices and updates its routing algorithm. Use of model based energy consumption of devices can reduce the amount of overhead from pulling the actual energy consumption information [12]. However the model must be also refreshed periodically to correct drift from actual energy usage. The field device can in addition notify the network manager when its remaining battery level drops below certain threshold so rerouting can be performed.

3.2.2 Link scheduling

While TDMA ensures collision free transmission, link assignment or scheduling must be carefully designed, as it impacts latency, reliability and energy efficiency. Within the WirelessHART superframe, the network manager decides which link to be activated for transmission or reception on which time slot. For two-way communication, two time slots are needed for each device pair. The link scheduling scheme should be simple (little overhead), scalable, and adaptive to for example different latency requirement of application data and make effective use of available bandwidth, so not many time slots are left unused in the superframe. Optimal time slot assignment, meaning assigning time slots to logical links using minimum number of timeslots [1] turns out to be an NP-hard problem [14]. However suboptimal algorithms can be applied for good performance. In [15], link scheduling was modeled as an *edge colouring* problem for a graph, where valid edge colouring is defined as no two edges incident on the same node are assigned the same colour. Consider the sensor network as a graph where each edge symbolizes logical link connection between two devices. The link scheduling problem can be solved mapping time slots to a colour in valid edge colouring. (more on the algorithm) The "hidden terminal"-like problem in TDMA schedule can be avoided by requiring each link operating on a different frequency, which is easily applicable in WirelessHART the PHY layer utilizes frequency hopping.

In [16], two centralized heuristic methods were proposed to solve the NP-hard TDMA link scheduling problem that is formulated specifically for WirelessHART like transmission setting, where many sensors or devices are to transmit data to the gateway. Due to the mesh network topology, spatial redundancy is possible in terms of link assignment. That is, having more than one node assigned to the same time slot at different places of the network where physical interference can be avoided. It is also required that no device is transmitting and receiving in the same time slot. Their objective is to determine the smallest length slot assignments that satisfy these constraints. In the node based heuristic scheduling, colouring (as proposed in [17]) is performed first on the given network topology. Then each device that has at least one packet to transmit at the beginning of the superframe will transmit at least one packet for the duration of the superframe. For the level based scheme, colouring is performed on a transform of the original network, which consists of different levels. Devices belong to the same level that are non-conflicting are scheduled first, followed by additional devices when possible. Simulation results showed that node based scheduling performs well when devices are transmitting packets are of relatively equal densities over the network, while the level based scheduling delivers better balanced packet transmission especially when more packets to be transmitted at lower levels of the network. Both schemes can be adopted by the WirelessHART protocol depending on the specific application requirement.

The link scheduling problem becomes more complex when there is an energy constraint. Energy waste can be caused by for example device operating in idle mode, or when buffer overflows at intermediate routing nodes in a multihop setting. In [18], an energy efficient link scheduling algorithm was proposed for TDMA MAC layer. Their optimization objective was to minimize the energy consumed from device idling and transition between active and sleep states, with the same conflict constraints stated above in [16]. BFS (breath first search) was used to generate the initial solution for the heuristic search method. The iterative tabu-search [19] is carried out on three different levels of network structure: tree, node and branch. Simulation were conducted to compare the proposed scheduling algorithm with DFS and BFS which are without energy constraint. Results showed a maintained low average energy consumption on the device level as the size of the network grows.

4 Cross-layer Design Issues

The OSI protocol modeled shown in Figure 3 displays a classic layered architecture. Here each protocol layer acts as an independent module with its dedicated functions and handles data packets from coming the layer above or below it. The layered structure is simple and robust and has been proven to function well in the wired network. However, in a wireless network, typical characteristics such as shared transmission medium, limited resources and lossy communication channels promoted the paradigm of *cross-layer design*. Essentially, it allows communications between the different protocol layers and the actual functions can be designed and optimized jointly. The main cross-layer design benefits include better efficiency, throughput, better allocation of resources, lower delay, less or more effective energy consumption [20]. Within the WirelessHART standard, as it is not specified how the protocols should be implemented, cross-layer design can be considered as an alternative way of implementation to the traditional approach.

In this section, we focus on cross-layer design of MAC and Network layer, with special considerations on energy consumption. PHY layer techniques such as link adaptation is not considered since the WirelessHART PHY layer complies to the strict 802.15.4 standard.

TDMA link scheduling algorithms that mentioned above, can be, for example, optimized together with routing, in the sense of allowed flow rates on each link and at the same time minimize the total network energy consumption as proposed in [21]. In the case of simple string topology, i.e. a multihop path from the sender to the gateway, it was concluded that single hop transmission are more efficient. Alternatively, device life time can also be considered as an optimization criterion. It can be ex-

[1] We use "link scheduling", "time slot assignment" and "link assignment" interchangeably.

pressed as the ratio of total battery power available and the average power spent. Optimization of the link schedules for network life time optimization becomes maximization of the *minimum device life time*. In this case for the string topology, the optimal strategy involves a combination of single-hop transmission of some packets and multihop of others. This is to ensure the device closest to the gateway does not have its battery depleted too quickly.

To take the above described scheme one step further, load balancing can be directly incorporated in the optimization process as suggested in [22]. Load balancing refers to the objective of avoiding hot spots in the network where nodes are quickly drained of battery from relaying for nodes that are further away from the gateway. Again it can be formulated as a multi-constraint convex optimization problem and solved using an iterative algorithm at the gateway. By applying simple greedy heuristic at each iteration, links are adaptively scheduled to determine an optimal routing, transmission power and rate in each time slot. It was shown through simulation that there is clear gain in network lifetime compare to static TDMA link scheduling algorithms over different network topologies. This is the result of a combined effort through multihop routing, frequency reuse and load balancing.

5 Conclusions and future outlook

The newly released WirelessHART standard offers many opportunities of effectively utilizing wireless communications in HART device based industrial applications. In this paper, we summarize and present the standard in a concise and understandable manner, followed by giving recommendations on how the standard can be implemented with existing methods and algorithms. We pay special attention to issues that are important to achieve a stable and successful operation of the WirelessHART network.

As the HART Communication Foundation and the standardization committee continue to revise and update the WirelessHART specifications, there are still a number of remaining issues. For example, mobility has not been addressed at all within the standard scope, although the simple case of an operator with a handheld device moving around the plant facilities to configure and access data from the field devices is a common application scenario. Associated problems such as interference from time varying wireless channels, effective handover as the operator move from one network/device to another, localization, and constant change in network topology should be considered. The situation can be further complicated as we move to the stage of having numerous autonomous mobile robots/agents deployed on unmanned facilities to perform monitoring and maintenance tasks. Coordination, integration and co-existence issues will be more prevalent.

Within the existing scope of WirelessHART, upgrades can be incorporated to improve the network performance. For example, currently the PHY layer of the network devices complies to the IEEE 802.15.4 standard, which is designed for low power, low data rate devices with low complexity. The WirelessHART gateway is defined as a type of network device with network access points to ensure connectivity and throughput to the field device networks. Enhancement can be made for example using multiple-input multiple-output (MIMO) operations on the PHY layer. Higher data rates and wider bandwidth channels are the two main benefits at the cost of certain degree of digital signal processing, which is not of major concern as the gateway is self-powered and is designed to have much higher processing abilities compare to the field devices. The upcoming IEEE 802.11.n standard incorporates MIMO technology on top of the existing 802.11 WLAN. Similar amendments, which are compatible with the existing WirelessHART PHY layer can be made to improve network throughput, which may be particularly useful when more bandwidth demanding applications and devices are integrated into the network.

Even with the existing gateways, the network lifetime can be extended through routing data to *different* virtual gateways/access points, as the standard allows multiple gateways within one network. It was shown in [23] that mobile gateways, together with the proper routing algorithm can well balance the load distribution in the network. In the case of jointly optimized routing and mobility trajectory, the network lifetime can be well extend to 5 times of that with a static gateway. Although WirelessHART gateways are assumed static, with a intelligent placement strategy, the network manager can well construct a similar schedule for each gateway access points to be utilized in an alternating manner. Proper formulation and optimization of joint gateway location and schedule can be an interesting problem to look into.

References

[1] HART Communications Foundation, *http://www.hartcomm.org/index.html.*

[2] IEEE Standardization WPAN Task Group, *http://www.ieee802.org/15.*

[3] ZigBee PRO Specification Oct. 2007, *http://www.zigbee.org*

[4] ISA100, *"ISA100.11a, Release 1 An Update on the Process Automation Applications Wireless Standard", ISA Seminar, Orlando, Florida, Feb. 2008*

[5] F. Sivrikaya and B. Yener, "Time synchronization in sensor networks: a survey," *IEEE Network*, vol. 3, no. 3, pp. 325–349, May 2005.

[6] J. Elson; L. Girod and D. Estrin, "Fine-grained time synchronization using reference broadasts," in *Proceedings of ACM Symposium on operating systems design and implementation*, 2002.

[7] S. Ganeriwal; R. Kumar and M. Srivastava, "Timing sync protocol for sensor networks," in *Proceedings of ACM Conference on Embedded Networked Sensor Systems SENSYS*, 2003.

[8] M. Sichitiu and C. Veerarittiphan, "Simple, accurate time synchronization for wireless sensor networks," in *Proceedings of IEEE Wireless Communications and Networking Conference WCNC*, 2003.

[9] V. Greunen and J. Rabaey, "Lightweight time synchronization for sensor networks," in *Proceedings of ACM International Conference on Wireless Sensor Networks and Applications*, 2003.

[10] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," *Elsevier Ad Hoc Networks*, vol. 3, no. 3, pp. 325–349, May 2005.

[11] D. Braginsky and D. Estrin, "Rumor routing algorithm for sensor networks," in *Proceedings of ACM Workshop on Sensor Networks and Applications*, 2002.

[12] M. A. Youssef; M. F. Younis and K.A. Arisha, "A constrained shortest-path energy-aware routing algorithm for wireless sensor networks," in *Proceedings of IEEE Wireless Communications and Networking Conference WCNC*, 2002.

[13] R. C. Shah and J. M. Rabaey, "Energy aware routing for low energy ad hoc sensor networks," in *Proceedings of IEEE Wireless Communications and Networking Conference WCNC*, 2002.

[14] E. L. Lloyd and S. Ramanathan, "On the complexity of link scheduling in multi-hop radio networks," in *Proceedings of Conference on Information Science and Systems*, 1992.

[15] S. Gandham; M. Dawande and R. Prakash, "Link scheduling in sensor networks: distributed edge coloring revisited," in *Proceedings of Joint Conference of the IEEE Computer and Communications Societies, INFOCOM*, 2005.

[16] S.C. Ergen and P. Varaiya, "Tdma scheduling algorithms for sensor networks," in *University of California, Berkely, Tecnial Report*, 2005.

[17] R. Ramaswami and K. K. Parhi, "Distributed scheduling of broadcasts in a radio network," in *Proceedings of Joint Conference of the IEEE Computer and Communications Societies, INFOCOM*, 1989.

[18] G. Jolly and M. Younis, "An energy-efficient, scalable and collision-free mac layer protocol for wireless sensor networks," *Wireless Communications and Mobile Computing*, vol. 5, no. 3, pp. 285–304, 2005.

[19] A. Dell'Amico and A. Trubian, "Applying tabu search to the job-shop scheduling problem," vol. 41, no. 3, pp. 231–252, sep 1993.

[20] A. Goldsmith and S. B. Wicker, "Design challenges for energy-constrained ad hoc wireless networks," *IEEE Wireless Communications Magazine*, vol. 9, no. 4, pp. 8–27, Aug. 2002.

[21] S. Cui; R. Madan; A. Goldsmith and S. Lall, "Energy-delay tradeoffs for data colleciton in tdma based sensor networks," in *Proceedings of the International Conference on Communications (ICC)*, 2005.

[22] R. Madan; S. Cui; S Lall and A. Goldsmith, "Cross-layer design for lifetime maximization in interference-limited wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 5, no. 11, pp. 3142–3152, Nov. 2006.

[23] J. Luo and J. Hubaux, "Joint mobility and routing for lifetime elongation in wireless sensor networks," in *Proceedings of Joint Conference of the IEEE Computer and Communications Societies, INFOCOM*, 2005.