Real-Time Wireless Sensor-Actuator Networks for Industrial Cyber-Physical Systems

Chenyang Lu, Senior Member, IEEE, Abusayeed Saifullah, Member, IEEE, Bo Li, Mo Sha, Member, IEEE, Humberto Gonzalez, Member, IEEE, Dolvara Gunatilaka, Chengjie Wu, Lanshun Nie, and Yixin Chen, Senior Member, IEEE.

Abstract—With recent adoption of wireless sensor-actuator networks (WSANs) in industrial automation, industrial wireless control systems have emerged as a frontier of cyber-physical systems. Despite their success in industrial monitoring applications, existing WSAN technologies face significant challenges in supporting control systems due to their lack of real-time performance and dynamic wireless conditions in industrial plants. This article reviews a series of recent advances in real-time WSANs for industrial control systems: (1) real-time scheduling algorithms and analyses for WSANs; (2) implementation and experimentation of industrial WSAN protocols; (3) cyber-physical co-design of wireless control systems that integrate wireless and control designs; (4) a wireless cyber-physical simulator for co-design and evaluation of wireless control systems. This article concludes by highlighting research directions in industrial cyber-physical systems.


I. INTRODUCTION

Wireless sensor-actuator networks (WSANs) are gaining rapid adoption in process industries due to their advantage in lowering deployment effort in harsh industrial environments. Industrial standard organizations such as ISA [1], HART [2], WINA [3] and ZigBee [4] have been actively pushing the application of wireless technologies in industrial automation and manufacturing. While early success of industrial WSANs focused on monitoring applications, there is significant value in exploring WSANs for process control applications to take full advantage of wireless technology in industrial plants.

However, wireless control systems face unique challenges that distinguish them from traditional control systems. First, it is challenging to meet the stringent latency requirements of feedback control in WSANs. IEEE 802.15.4 radios commonly used in WSANs have a maximum bandwidth of only 250 kbps. Multi-hop communication over mesh networks further increases communication delays. Furthermore, channel conditions in industrial environments can change dynamically due to external interference, moving obstacles and weather conditions. As a result, WSANs demand real-time scheduling algorithms and analysis techniques that are both effective and efficient. In contrast to wired industrial networks (e.g., Control Area networks [5]) with well established real-time scheduling techniques, there have been limited results on real-time scheduling for WSANs. The lack of analytical methods for achieving real-time performance in WSANs hinders their adoption in control systems. Wireless control demands a new real-time scheduling theory for WSANs [6].

Second, design of wireless control systems must deal with interdependencies between control and communication. For example, while it is well known in digital control that a low sampling rate usually degrades control performance [7], a high sampling rate may increase resource contention in bandwidth-constrained WSANs leading to long communication delays, which again may lead to degraded control performance [8]. The coupling between wireless communication and control therefore motivates a cyber-physical co-design approach that integrates wireless networks and control designs.

This article provides a review of recent advances in real-time WSANs for industrial control systems, an increasingly important class of cyber-physical systems (CPS) in the dawn of Industrial Internet [9] and Industry 4.0 [10].

Section II describes real-time scheduling for industrial WSANs. We first introduce WirelessHART, an industrial standard for wireless sensing and actuation. We then present a suite of real-time scheduling algorithms, delay analyses and protocol implementation to achieve real-time performance in WirelessHART networks. Section III introduces our recent efforts to design industrial wireless control systems following a cyber-physical co-design approach. We first study the problem of selecting sampling rates to optimize control performance. We then investigate how to incorporate emergency alarms in wireless control systems. To support the design and evaluation of wireless control systems, we introduce the Wireless Cyber-Physical Simulator (WCPS) for holistic studies of wireless control systems. We conclude this article by highlighting promising research directions.

II. REAL-TIME INDUSTRIAL WSANS

We first provide an overview of the WirelessHART standard and then describe our implementation and experimentation of a
WirelessHART protocol stack. Finally, we summarize our real-time scheduling algorithms and analyses for WirelessHART.

A. WirelessHART

A wireless control system comprises feedback control loops connecting sensors, controllers and actuators through a wireless mesh network. Sensors measure variables of the plant and send the measurements to a controller over the wireless mesh network. The controller then sends control commands to the actuators in order to control the physical processes. Industry plants pose harsh environments for wireless communication due to significant channel noise, physical obstacles, multi-path fading, and interference from coexisting wireless devices. These harsh and dynamic environments make it difficult for WSANs to meet the stringent reliability and real-time requirements of industrial control applications.

There exist multiple industrial WSAN standards, e.g., WirelessHART, ISA100, WINA and ZigBee. In this article we will focus on the WirelessHART standard, although the design principles may be extended for the other standard technologies. WirelessHART devices are now adopted worldwide for industrial process management and control. Unique features that make WirelessHART particularly suitable for industrial process control are described below.

A WirelessHART network is managed by a centralized network manager. The network manager is responsible for the management, scheduling, creating the routes, and optimizing the network. Network devices include a Gateway, a set of field devices, and several access points as shown in Fig. 1. The network manager and the controllers are installed or connected to the Gateway. The field devices are wireless sensors and actuators. Each field device is equipped with a half-duplex omnidirectional radio transceiver compliant with the IEEE 802.15.4 standard. Multiple access points are wired to the Gateway to provide redundant paths between the wireless network and the Gateway.

A WirelessHART network adopts various mechanisms to ensure reliable communication in unreliable industrial environments. Time is globally synchronized and slotted using a built-in time synchronization protocol. Each time slot is 10 ms, which is sufficient to send or receive one packet and an corresponding acknowledgement. Transmissions are scheduled based on a multi-channel TDMA (Time Division Multiple Access) protocol. A time slot can be either dedicated or shared. In a dedicated slot, only one sender is allowed to transmit to the receiver. In a shared slot, more than one sender can attempt to transmit to the same receiver. Since shared slots are assigned to multiple senders, collisions may occur within a shared slot. To reduce collisions, senders adopt Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) to contend in a shared slot. CSMA/CA allows multiple nodes to share a common channel through carrier sensing, where nodes attempt to avoid collisions by transmitting only when the channel is idle.

The network uses 16 channels defined in IEEE 802.15.4, and adopts channel hopping in every time slot. Channel hopping provides frequency diversity to mitigate interferences and reduce multi-path fading effects. Any excessively noisy channel is blacklisted.

A WirelessHART network supports two types of routing approaches: source routing and graph routing. Source routing provides a single directed path for routing from a source to a destination device. Graph routing involves a routing graph consisting of a directed list of paths between the two devices, and is adopted for enhanced end-to-end reliability. In graph routing, packets from field devices are routed to the Gateway through the uplink graph. To every field device, there is a downlink graph from the Gateway. The end-to-end communication between a source (sensor) and destination (actuator) happens in two phases. In the sensing phase, on one path from the source to the Gateway in the uplink graph, the scheduler allocates a dedicated slot for each device starting from the source, followed by allocating a second dedicated slot on the same path to handle a retransmission. Then, to offset failure of both transmissions along a primary link, the scheduler allocates a third shared slot on a separate path to handle another retry. Then, in the control phase, using the same method, the dedicated links and shared links are scheduled in the downlink graph of the destination.

B. Implementation and Experimentation

To study and evaluate WSAN protocols, we have developed an experimental WSAN testbed. The system comprises a network manager on a server and a network protocol stack implementation on TinyOS 2.1.2 and TelosB motes. Each mote is equipped with a TI MSP430 microcontroller and a TI CC2420 radio compatible with the IEEE 802.15.4 standard. Figure 2 shows the motes deployment in our campus buildings. In an experiment, our motes can be designated as access points and field devices, which form a multi-hop wireless mesh network running WSAN protocols. To support experimentation and measurements, all motes in our system are physically connected to a wired backplane network that can be used for managing wireless experiments and measurements without interfering with wireless communication.

Our network manager implements a route generator and a schedule generator. The route generator is responsible for generating source routes or graph routes based on the collected
We have performed a series of experiments on our WSAN testbed [14]. For instance, we have conducted a comparative study of the two alternative routing approaches adopted by WirelessHART, namely source routing and graph routing, and investigated the tradeoff among reliability, latency, and energy consumption under the different routing approaches.

We first run our experiments in a clean environment where we blacklist the four 802.15.4 channels overlapping with our campus Wi-Fi network and run the experiments on the remaining 802.15.4 channels. We then repeat our experiments in a noisy environment where we configure the network to use channels overlapping with our campus Wi-Fi network and in a stress-testing environment where we generate controlled Wi-Fi interference. In our study, we use reliable links with PRR higher than 80%, where PRR stands for Packet Reception Ratio, i.e., the fraction of transmissions that are successful over a wireless link. We set up 8 data flows, and run our experiments long enough such that each flow delivers at least 500 packets from its source to its destination. We measure the packet delivery rate (PDR) of a flow, defined as the percentage of packets that are successfully delivered to their destination. Table I shows the minimum, median and maximum PDR among all the flows in the clean, noisy and stress-testing environments. Graph routing improves the median PDR by 15.9%, and 21.4% compared to source routing in noisy and stress testing. Moreover, graph routing drastically improves the minimum PDR, with 35.5% and 63.5% increase in noisy and stress testing environments. This result shows graph routing can effectively enhance the predictability of real-time performance, especially under significant interference. However, our study also shows that route diversity incurs costs for latency and energy consumption, with graph routing suffering from an average of 80% increase in end-to-end latency and consuming an average of 130% more energy over source routing.

Our study concludes that it is important to employ graph routing algorithms specifically designed to optimize latency and energy efficiency. This problem was recently explored in [20], which proposed a set of real-time and energy-efficient routing algorithms for WirelessHART networks. Both source routing and graph routing have been implemented in our WSAN testbed and the Wireless Cyber-physical Simulator (WCPS) (to be discussed in III.B), which enable systematic comparison between the alternative routing approaches.

C. Real-Time Scheduling

To save cost and enhance flexibility, it is desirable to support many control loops in a same network. The stability and control performance of these systems heavily depends on real-time communication over the shared wireless network. Feedback control loops in a WSAN therefore impose strict requirements on reliability and real-time guarantees in wireless communication in order to avoid plant shutdowns and/or accidents. For example, in oil refineries, spilling of oil tanks has to be avoided by controlling oil level in real-time. However, industrial plants poses a harsh environment for wireless communication due to unpredictable channel conditions, limited bandwidth, physical obstacles, multi-path fading, and interference from coexisting wireless devices [11]. With the adoption of industrial wireless standards such as WirelessHART [2], process monitoring and control applications have seen the feasibility of achieving reliability and real-time wireless communication through spatial and spectrum diversity. Real-time communication in these wireless networks pose new and important challenges. Unlike wired networks, there have been limited results on real-time scheduling theory for wireless networks. Real-time transmission scheduling and

<table>
<thead>
<tr>
<th>Environment</th>
<th>Source Routing</th>
<th>Graph Routing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min</td>
<td>Med</td>
</tr>
<tr>
<td>Clean</td>
<td>0.97</td>
<td>0.99</td>
</tr>
<tr>
<td>Noisy</td>
<td>0.65</td>
<td>0.82</td>
</tr>
<tr>
<td>Stress Testing</td>
<td>0.29</td>
<td>0.70</td>
</tr>
</tbody>
</table>
analysis for WSANs require new methodologies to deal with unique wireless characteristics.

1) Overview of State of the Art: Real-time communication in wireless networks has been explored in many earlier efforts [21]–[32], [32]–[38]. The survey in [39] provides a comprehensive review of these works. However, these are not suitable for industrial applications that usually need multi-channel communication, multi-path routing and real-time performance analysis results. Earlier works [40]–[44] on real-time performance analysis for wireless sensor networks focused on data collection through a routing tree [40], [43].

Real-time WSAN for control purposes has been studied for single hop networks in [45]–[49]. For multihop wireless networks, a mathematical framework to model and analyze schedules for WirelessHART networks has been proposed in [50]. Real-time scheduling for WirelessHART networks has received attention in recent works [51]–[57]. These scheduling policies can be broadly classified into two categories: fixed priority scheduling and dynamic priority scheduling [58]. In a fixed priority scheduling policy, each data flow has a fixed priority, and every transmission has the priority of its flow. Due to its simplicity and low scheduling overhead, fixed priority scheduling is a common class of real-time scheduling policies in practice [58]. On the other hand, dynamic priority scheduling policy refers to the policy where there is no fixed priority i.e., priorities of transmissions during the scheduling change dynamically according to some chosen criterion. While such policies have higher scheduling overhead, they provide better schedulability since priorities are changed dynamically to enhance the schedulability.

In the following we summarize our recent results on real-time scheduling for industrial WSANs under fixed priority and dynamic priority scheduling, respectively.

2) Fixed Priority Scheduling: Delay analysis can be used to determine whether real-time data flows can meet their deadlines. When used offline, the analysis helps users design a WSAN to meet the real-time performance requirements. For example, we used schedulability analysis to select the sampling rates of wireless control loops to optimize control performance under communication deadlines [59]. In [54], we used schedulability analysis to assign priorities to real-time flows in order to meet their deadlines. When used online, the network manager can adjust the workloads in response to network dynamics. Industrial environments can cause frequent changes in network topologies and channel conditions. If a WSAN can no longer guarantee the deadlines for all the flows based on delay analysis, the network manager may stop or adjust the data rates of non-critical flows in order to maintain real-time guarantees to the remaining flows. The Wireless Cyber-Physical Simulator (WCPS) introduced in Section III-B can be used to assess the pessimism of delay analysis and wireless control performance under realistic settings.

For fixed priority scheduling in WirelessHART networks, we proposed a suite of delay analyses for real-time flows [52], [53]. These analyses determine upper bounds on the end-to-end communication delays and provide sufficient conditions for schedulability. Two factors contribute to the communication delay of a control loop. A lower priority flow can be delayed by higher priority flows (a) due to channel contention (when all channels are assigned to transmissions of higher priority flows in a time slot), and (b) due to transmission conflicts (when a transmission of the flow and a transmission of a higher priority flow involve a common node). A key insight underlying this analysis is to map the real-time transmission scheduling in WirelessHART networks to real-time multiprocessor scheduling. By incorporating the unique characteristics of WirelessHART networks into the state-of-the-art worst case response time analysis for multiprocessor scheduling, the analysis can efficiently compute a safe and tight upper bound of the end-to-end delay of every flow.

Priority assignment is an important problem in fixed-priority scheduling as it has a significant impact on the schedulability of real-time flows. An ideal priority assignment should not only enable real-time flows to meet their deadlines, but also work synergistically with real-time schedulability tests to support effective network capacity planning and efficient online admission control and adaptation. In [54] we studied optimal and near optimal fixed priority assignment using local search. A salient feature of this local search approach is that it exploits the delay bounds provided by a schedulability analysis to efficiently search for feasible priority assignments by discarding unnecessary branches in the search space.

3) Dynamic Priority Scheduling: Dynamic priority scheduling of transmissions was studied in [55], [57] for WirelessHART networks of with tree topologies. Our work in [51], [56] studied dynamic priority scheduling of transmissions for general WirelessHART networks. The work in [51] has shown that the real-time transmission scheduling for WirelessHART networks is NP-hard. We observed that transmission conflicts play a major role in the communication delays and the schedulability of the control loops, which makes the traditional real-time scheduling policies such as Least Laxity First (LLF) less effective. We propose an optimal local search scheduling algorithm that exploits the necessary condition to effectively discard infeasible branches in the search space. We also proposed a faster heuristic called Conflict-aware Least Laxity First (C-LLF) for dynamic priority scheduling. The algorithm identifies the critical time windows in which too many conflicting transmissions have to be scheduled, thereby determining the criticality of each transmission. Criticality of a transmission is quantified by its conflict-aware laxity, which is its laxity after discarding the (estimated) time slots that can be wasted through waiting to avoid transmission conflict. Transmissions exhibiting the lowest conflict-aware laxity are assessed to be more critical. C-LLF gives the highest priority to the transmissions exhibiting lower conflict-aware laxity. Thus C-LLF integrates LLF and the degree of conflicts associated with a transmission, and outperforms traditional real-time scheduling policies.

Our work in [56] has provided a schedulability analysis for earliest deadline first (EDF), a common dynamic priority scheduling in WirelessHART networks.

We evaluated our schedulability analyses against experimental results on our WSAN testbed [17] as well as in simulations. In our evaluation, both fixed-priority scheduling [54] and
dynamic priority scheduling [51] policies outperformed the traditional real-time scheduling policies by significant margins. All experimental results and simulations showed that our analyses provide safe upper bounds of communication delays in the network and enable effective schedulability tests [53]. For tighter upper bounds under graph routing, we established probabilistic delay bounds in [6] that represent upper bounds with probability $\geq 0.90$. The worst-case and probabilistic bounds can be used in different application scenarios depending on the level of predictability required. Our analysis hence can be used for an effective schedulability test and admission control of real-time flows in WirelessHART networks.

III. WIRELESS CONTROL CO-DESIGN

Due to limited resources in a WSAN shared by multiple control loops, it is critical to optimize the overall control performance. However, in a wireless control system, the control performance not only depends on the control algorithms, but also relies on real-time communication over the shared wireless network. The coupling between real-time communication and control requires a cyber-physical co-design approach for a holistic optimization of control performance. This section summarizes our recent efforts on cyber-physical co-design for industrial wireless control systems. We first study the problem of selecting sampling rates to optimize the control performance of multiple feedback control loops sharing a WSAN. To support cyber-physical co-design of wireless control systems, we present Wireless Cyber-Physical Simulator (WCPS), an integrated simulation environment for holistic studies of wireless control systems. We then investigate how to incorporate emergency alarms in wireless process control systems. We wrap up this section with a case study of wireless process control for coupled water tanks using both regular and emergency control loops.

A. Rate Selection

In control systems, the impacts of sampling rate have been studied for robot control [60], disturbance rejection [61], and various control performance indices [62]–[66]. The choice of sampling rates needs to balance control and communication. While a high sampling rate is desirable from a pure digital control perspective, it also has the undesirable effect of heavier network load and longer communication delay, which can also degrade the control performance.

The control performance can be quantified as a function of sampling rates using the formulation proposed by Seto et al. [62], which characterizes the performance difference between the continuous-time control system and that of its digital implementation.

In [59], [67] we apply these ideas to optimize the performance of a set of feedback control loops over a WSAN, where the control cost of the $i$-th control loop at the sampling rate $f_i$ is approximated as $\alpha_i e^{-\beta_i f_i}$, where $\alpha_i, \beta_i > 0$ are the magnitude and decay rate coefficients, respectively. Given the sampling rates of all of the control loops $\{f_1, f_2, \ldots, f_n\}$, the total control cost of the system is defined by:

$$\sum_{i=1}^{n} w_i \alpha_i e^{-\beta_i f_i},$$  \hspace{1cm} (1)

where $w_i$ is the weight of the $i$-th control loop. We then use the total control cost as the performance index of the system and as the objective of an optimization problem to determine the best sampling rates of a multi-loop control system sharing a WirelessHART network under stability and delay constraints. In particular, the objective is to minimize the cost in (1) subject to two constraints. First, the delay bound $R_i$ of the $i$-th loop, which can be computed based on the sampling rate in [52], [53], must be smaller than its predetermined deadline $D_i$. Second, each control loop must have a sampling rate $f_i$ within its minimum and maximum rates, denoted $f_i^{\text{min}}$ and $f_i^{\text{max}}$ respectively, to ensure stability. Thus, we mathematically formulate this optimization problem as follows:

$$\min_{\{f_i\}_{i=1}^{n}} \sum_{i=1}^{n} w_i \alpha_i e^{-\beta_i f_i},$$  \hspace{1cm} (2)

subject to: $R_i(f_1, \ldots, f_n) \leq D_i, \forall i \in \{1, \ldots, n\},$ $f_i^{\text{min}} \leq f_i \leq f_i^{\text{max}}, \forall i \in \{1, \ldots, n\}.$

The resulting constrained optimization problem is challenging since the delay $R_i$ is non-differentiable, non-convex, and not in closed-form. We explored four methods to solve this problem: (a) a subgradient-descent algorithm, (b) a greedy heuristic, (c) a penalty method using simulated annealing (SA), and (d) a convex relaxation method based on a new delay bound that is convex and smooth but more pessimistic than previous analyses [52], [53] using a different approach. Notably, the convex relaxation greatly simplifies the optimization problem, since the overall problem becomes convex and differentiable by simplifying the delay bound analysis.

In [59], we evaluate the methods through simulations based on the topology of our WSAN testbed [17]. The results demonstrate that SA achieves the lowest control cost but requires the longest execution time. Both the greedy heuristic and the subgradient method lead to high control costs because the optimization problem is highly nonlinear with a large number of local extrema. For example, for 30 control loops, the greedy heuristic incurs a control cost up to 2.67 times that of SA. In contrast, the convex relaxation approach using an interior point method incurs a control cost no more than 35% higher than that of SA, while incurring a much shorter execution time. The convex relaxation approach therefore achieves the desirable balance between control performance and run-time efficiency. This result demonstrates the significant advantage of cyber-physical co-design that integrates control optimization and schedulability analysis in wireless control systems.

B. WCPS

Existing wireless control system research often relies on lab-scale equipment and simulations. However, lab-scale equipment usually suffers from limited physical size, which cannot capture delays and data loss in real WSANs. Simulation tools
for control systems often lack realistic models of WSANs that exhibit complex and stochastic behavior in real-world environments. The lack of tools that can capture both the cyber wireless network and physical aspects of control systems has been a hurdle to wireless control research.

Early experimental work on wireless control [68]–[73] usually relied on a lab testbed where wireless sensors are within a single hop and experience no data loss due to the physical proximity of the devices. The challenge in realistic experimentation with wireless control systems motivated the development of simulation tools for wireless control systems. For example, NCSWT [74] and Gisso [75] are two simulators designed for wireless control systems. TrueTime [7] is a well-established control system simulator capable of holistic studies of CPU scheduling, communication, and control algorithms. Unfortunately, none of these simulators provides a realistic wireless radio model or a state-of-the-art WSAN protocol stack.

Wireless Cyber-Physical Simulator (WCPS) [76] is designed to provide a holistic and realistic simulation of wireless control systems. WCPS employs a federated architecture that integrates (a) Simulink for simulating the physical system dynamics and controllers, and (b) TOSSIM for simulating WSANs. Simulink is commonly used by control engineers to design and study control systems, while TOSSIM [77] has been widely used in the sensor network community to simulate WSANs based on wireless link models that have been validated in diverse real-world environments [78]. WCPS provides an open-source middleware to orchestrate simulations in Simulink and in TOSSIM.

WCPS 2.0 implements the WirelessHART network protocol stack at the routing and MAC layers [79]. To support WirelessHART we also extended TOSSIM to simulate wireless communication over multiple channels. We have implemented both Source Routing and Graph Routing as specified in the WirelessHART standard. To our knowledge, WCPS 2.0 is the first simulator that supports WirelessHART protocols based on a realistic wireless link model.

Thanks to WCPS we have been able to develop case studies that simulate wireless control systems for civil infrastructure and process plants. In [80] we simulated two wireless structural control systems. Wireless structural control systems offer an attractive approach to protect civil infrastructures from natural hazards such as earthquakes and other natural disasters.

In the first simulation we studied a benchmark building model, where the wireless traces were collected in a multi-story building. In the second simulation we studied the structural model of the Cape Girardeau bridge over the Mississippi River, where the wireless traces were collected from a similar bridge in South Korea. These case studies shed light on the limitations of traditional structural control approaches under realistic wireless conditions. They further allowed us to validate the advantages of cyber-physical co-design algorithms for wireless communication protocols. Based on our experience with WCPS, we have recently enhanced the wireless building control study, transforming it into a benchmark that can be used by the structural control community to explore and evaluate different wireless control approaches [81], allowing practitioners to easily generate and configure realistic nuisances such as network induced delay, data loss, measurement noise, and control constraints.

C. Emergency Communication

In [79] we considered the case of WirelessHART control networks comprised of (a) regular flows, which are periodically generated and typically stabilize a desired part of a physical system, and (b) emergency flows, which are infrequent and typically signal that an unsafe situation is about to occur. Both regular and emergency flows have predetermined deadlines to transmit a packet. Since emergency flows carry information related to unsafe situations, these flows have a higher criticality than regular flows. Therefore, we proposed the following approach to transmit packets through the network: (a) in regular mode, i.e., when there is no emergency, all regular flows should meet their deadlines, and (b) in emergency mode, i.e., when an emergency occurs, all existing emergency flows should meet their deadlines, while regular flows are delivered on a best-effort basis.

Periodic Scheduling (PS) is a baseline scheduling approach that reserves periodic slots dedicated to emergency packets in the transmission schedule. Emergency packets can be transmitted only in the reserved slots. When there is no emergency, the reserved slots are left unused. We take a two-level priority assignment approach for scheduling, where we first schedule emergency flows to meet their deadlines, and then we schedule the regular flows using the remaining capacity. Within each group, flows are scheduled using a rate monotonic policy. This basic approach is simple in design, yet it wastes (a potentially large amount of) network bandwidth, reserving time slots for emergency flows even when there is no emergency.

To avoid wasting resources when no emergencies occur, we designed a Slot Stealing (SS) method that allows emergency flows to “steal” slots from regular schedules when emergencies arise, removing the need to allocate exclusive time slots for emergency flows. The “stealing” process is implemented by adding a random backoff and a Clear Channel Assessment (CCA) to the transmission of all regular flow packets, and allowing emergency packets to be transmitted immediately at the beginning of the time slot. Hence, if an emergency packet is available, then it will take the slot, and any regular flow packet trying to use that slot will fail the CCA, forcing it to wait for the next available time slot or drop the packet. By allowing the overlapping emergency and regular schedules, SS is able to schedule the same number of flows in a shorter time frame when compared to PS, while ensuring timely delivery of emergency packets in case of an emergency.

We also explored two alternative approaches to send emergency signals. For systems that need to periodically monitor and control the emergency state, an emergency flow is activated whenever emergencies occur, and the emergency flow then periodically generates data until the emergency is over. For systems that do not need to periodically monitor and control the emergency state, the system can adopt an event-based approach to communicate the emergency alarms, i.e., an emergency sensor only sends one alarm-start packet and one
alarm-stop packet at the beginning and end of the emergency, respectively. While the event-based communication results in the same transmission schedule, it significantly reduces the number of regular transmissions that are dropped or delayed by emergency transmissions, leading to better control performance. In [79] simulation results showed the combination of slot stealing and event-based emergency communication produced the best results in terms of emergency handling and control performance.

**D. Emergency Control Case Study**

We tested our emergency communication protocol on a set of two identical water tank systems sharing a common wireless network simulated using WCPS 2.0. A diagram of each of water tank systems is shown in Fig. 4. Our choice is based on the simple, yet representative, dynamics of water tanks, their hybrid dynamical nature (since the evolution of the system changes when the water tanks are either full or empty), and more importantly, its similarity to systems commonly used in industrial applications. Also, we simulated two identical water tank systems sharing a wireless network to study the effect of one system’s emergency over the second system’s regular flow.

Following the software architecture in WCPS 2.0, the sensor data generated by Simulink is fed into the WSAN simulated using TOSSIM. TOSSIM then returns the packets delayed or dropped according to the behavior of the network, which are then fed to the controller implemented in Simulink. Controller commands are then fed again into TOSSIM, which delays or drops the packets and sends the outputs to the actuators in the water tanks, closing the loop.

For this study we collect wireless traces from 21 nodes in a WSAN testbed [17]. We then use the wireless traces as inputs to the TOSSIM simulator to generate realistic wireless characteristics in a simulated WSAN. The route in the simulated WSAN is 6 hops. We used further adjusted Received Signal Strength (RSS) to test the system performance with different wireless signal strengths.

On the control side, we used a PID controller to regulate the water level of Tank 2, denoted $L_2$, by actuating the Pump taking water from the Basin, as shown in Fig. 4. If no emergencies occur, then the Valve stays closed. We also defined a discrete emergency controller setting the values of the Pump and Valve, with the objective of avoiding water spillage. Hence, if any of the emergency level sensors, denoted $L_1^H$, $L_2^H$, $L_3^H$, and $L_4^H$, detect a dangerous water level, then an emergency signal is sent and the controller switches from the PID controller to the discrete emergency controller, switching back when the emergency is cleared. We defined a system failure using two principles: (a) if the system cannot stabilize at a regular configuration (i.e., within non-emergency water levels) in a fixed 100 second interval, and (b) if any of the water levels exceed the height of its associated water tank.

Our simulations show that the combination of the slot-stealing and event-based emergency communication is highly effective in avoiding system failures, since it produces a tighter schedule that results in a faster update frequency of the regular control flows. It also reduces communication load and mitigates the impact of emergency communication on regular control flows. Our case study demonstrates that even for a 6-hop lossy wireless network (with 5.8% median packet loss), successful system control can still be achieved through a combination of regular and emergency control.

**IV. RESEARCH DIRECTIONS**

**A. Enhancing Scalability**

A major limitation of current industrial WSANs is their limited scalability due to the centralized network architecture. In WirelessHART, when a node or link fails, the centralized network manager must generate a new global TDMA schedule, which requires the network manager to collect the current topology of the network, create a new schedule, and distribute the schedule among all field devices. As WSANs are subject to frequent changes in channel quality and link condition, global changes to the network schedule create excessive communication overhead in a large network. As a result, while the centralized architecture has proven to be sufficient for small-scale installations, it can become a significant limitation as WSANs start to be deployed over large geographic areas (e.g., thousands of devices over an oil field). A key research direction is to make WSANs scalable while achieving end-to-end real-time communication.

A promising approach to enhance scalability is through a hierarchical network architecture, where a large WSAN is divided into multiple subnetworks. Each subnetwork employs its own manager to manage local operations, and a global manager coordinates with the subnetwork managers to manage the entire network in a hierarchical fashion. An advantage of this architecture is that it can meet industrial needs for both network visibility and scalability. As a submanager may deal with the wireless dynamics within its subnetwork, the hierarchical architecture can scale effectively. A challenge in designing a hierarchical architecture is to deal with the interdependencies among the subnetworks that share the wireless spectrum and need to support flows traversing multiple subnetworks subject to end-to-end deadlines.
B. Exploring White Spaces

Another limiting factor of today’s WSANs stems from the short communication range of the IEEE 802.15.4 radios adopted by industrial standards such as WirelessHART. To overcome the short communication range, many WSANs form multi-hop mesh networks resulting in long communication delays and limited scalability. An opportunity to enhance the scalability of real-time WSANs arises from the opening of a new spectrum resulting from the transition to digital TV broadcasting globally and freeing up the VHF/UHF spectrum. White spaces refer to the allocated but locally unused TV spectrum. Since TV transmissions are in lower frequencies, white space transmissions have excellent propagation characteristics over long distance. They can easily penetrate walls and other obstacles, and hence hold enormous potential for industrial applications that need real-time communication over large geographic areas. Thanks to its long communication range and wall-penetration capability, a white space network will have small hop counts and will drastically reduce communication delays and protocol complexities.

While white spaces are mostly being tapped into for wireless broadband access to date, they also open up the opportunity to support highly scalable real-time industrial applications that have been challenging under existing WSAN technologies [82]. The characteristics and the application demands in industrial sensing and actuation pose unique challenges in adopting white spaces for industrial WSANs. Instead of high-throughput traffic, a WSAN should exploit the wide spectrum in white spaces to support low-data-rate communication from numerous field devices. Furthermore, a WSAN over white spaces must achieve high degrees of energy efficiency that is comparable to today’s WSAN technologies. It also needs to handle changes and variations in spectrum availability in white spaces while maintaining desired real-time performance.

C. Cyber-Physical Co-Design

While earlier work has shown the promise of cyber-physical co-design in wireless control systems through point designs and case studies, there exist opportunities to develop a theory and practice of cyber-physical co-design through a broader exploration of the interaction between the network protocol stack and control design. For example, if a controller employs a state observer to estimate system states based on intermittent observations, the controller can tolerate a certain degree of data loss from the sensors. A WSAN can exploit the controller’s resiliency against sensor data loss by reducing the route redundancy from sensors to controllers, while dedicating more network resources to enhance the reliability of communication to actuators. That is, the allocation of network resources should be dependent on the control design. Conversely, an actuator may buffer the control inputs from a model predictive controller and use previously received control inputs when the network fails to deliver the new control inputs. When the wireless condition deteriorates causing more data loss, the controller may increase its sampling rates or control horizon to increase its tolerance to data loss to the actuators. That is, the controller should adapt to network conditions. Note that the adaptation of network resources and controller sampling rates can be designed using approximated mathematical models, or using distributed extensions of the data-driven methods in [83]. Therefore, a wireless control system will need to co-design wireless networks and control in an interdependent fashion. To establish a unified cyber-physical co-design approach to industrial wireless control, it will be crucial to develop interfaces between the WSAN and the control components to maintain optimal control performance by adapting to changing network and physical dynamics in the wireless control systems. We need to further develop a suite of algorithms and analytical techniques to assure the safety of the control system under dynamic conditions. Finally, the research community will need long-term collaboration among computer science, control theory and domain experts for this inherently interdisciplinary research.

D. Cyber-Physical Testbed

The advancement and adoption of new WSAN technologies have been hindered by the lack of testbeds under realistic industrial settings. While cyber-physical simulators such as WCPS are useful for studying wireless control systems, they require realistic industrial plant models and wireless traces collected from real-world industrial settings. Despite best efforts on modeling techniques, simulations cannot replace physical components and networks with complex and sometimes unexpected dynamics. Unfortunately, full-scale industrial plants are often difficult to replicate for research due to space and safety constraints. A step beyond simulations might be hybrid testing technology that combines physical components and simulations for hardware-in-the-loop experiments. For example, a physical WSAN testbed may be carefully integrated with simulations of an industrial plant to study the impact of real wireless dynamics on control systems. Conversely, an industrial plant may be integrated with a WSAN simulator to evaluate potential wireless control designs without a physical WSAN deployment. This would require extending tools such as WCPS to support hardware-in-the-loop testing where both the WSAN and the plant can be replaced by physical implementations. Close partnership between industry and researchers will be critical to develop realistic cyber-physical testbeds for industrial wireless control systems.

V. Conclusion

Real-time WSANs are poised to play a key role in industrial automation in the era of Industry 4.0 [10] and Industrial Internet [9]. Recent research and industrial developments have demonstrated the promise of supporting real-time communication and control over WSANs. This article reviews recent advances in two related fronts: (a) real-time scheduling and analytical techniques for achieving real-time performance in industrial WSANs and (b) cyber-physical co-design of wireless control systems. We highlight the significant challenges and opportunities in cyber-physical systems research that crosscut wireless and control domains. Interdisciplinary collaboration and partnership among wireless and control researchers and industry communities will be crucial in realizing the vision of wireless industrial control.
ACKNOWLEDGMENT

This work was supported, in part, by US National Science Foundation through grants 1320921 (NeTS), 1035773 (CPS), 1017701 (NeTS) and 1144552 (NeTS) and National Natural Science Foundation of China (NSFC) through grant 61273038.

REFERENCES

differentiation in stateless wireless ad hoc networks,” in INFOCOM,
vol. 2, no. 2, pp. 457–466.
synchronization protocol,” in Proceedings of the 2nd International
multihop scheduling and medium access with delay and throughputs constraints,” in MobiCom, 2001.
[18] K. Karenos and V. Kalogeraki, “Real-time traffic management in sensor
Humberto Gonzalez (S’03-M’12) received the B.S. and M.S. degrees in electrical engineering from Universidad de Chile, Santiago, Chile, in 2005, and the Ph.D. degree in electrical engineering and computer sciences from the University of California at Berkeley, CA, USA, in 2012. He is currently an Assistant Professor in the Department of Electrical & Systems Engineering at Washington University in St. Louis. His research focuses on the theory and implementation of control algorithms for cyber-physical systems, with an emphasis on optimal control and numerical algorithms.

Dolvara Gunatilaka is a PhD student in the Department of Computer Science and Engineering at Washington University in St. Louis. She received her M.S. degree in Computer Science from the same school in 2013. Prior to her Master degree, she received her B.S. Degree in Information and Communication Technology from Mahidol University, Thailand in 2010. Her research focuses on industrial wireless sensor-actuator networks.

Chengjie Wu is a software engineer at Yahoo!. He received the Ph.D. degree in Computer Science from Washington University in St. Louis in 2014, the M.S. and B.S. degrees from Tsinghua University in 2008 and 2006, respectively. He was a recipient of the McDonnell International Scholarship and Emerson Corporate Fellowship awarded by Washington University in St. Louis. He received the Best Paper Nomination at the RTAS’12. His research interests include real-time systems, cyber-physical systems, wireless sensor networks, and Internet of Things.

Lanshun Nie received his B.S., M.S. and Ph.D degrees from the Harbin Institute of Technology, China. He was a visiting scholar of the CPS laboratory at Washington University in St. Louis. Lanshun is currently an Associate Professor of Computer Science at the School of Computer Science and Technology, Harbin Institute of Technology, China. His research interests involve Cyber-Physical Systems and Wireless Sensor Networks.

Yixin Chen is an Associate Professor of Computer Science at the Washington University in St. Louis. His research interests include data mining, machine learning, artificial intelligence, and optimization. He received the Best Student Paper Runner-up Award at the KDD’14, and Best Paper Awards at AAAI’10 and ICTAI’05. He has won First Prizes in the International Planning Competitions (2004 & 2006). He received an Early Career Principal Investigator Award from the Department of Energy (2006) and a Microsoft Research New Faculty Fellowship (2007). He is an Associate Editor for ACM Transactions of Intelligent Systems and Technology and serves on the Editorial Board of Journal of Artificial Intelligence Research.