Combining Admission and Modulation Decisions for Wireless Embedded Systems

John Meier, Christopher Gill, and Roger D. Chamberlain
Dept. of Computer Science and Engineering, Washington University in St. Louis
One Brookings Drive, St. Louis, Missouri 63130
Email: {jmeier,cdgill,roger}@wustl.edu

***Abstract*—Wireless communication is increasingly being used to federate embedded devices in a variety of distributed systems application domains, ranging from wireless sensor networks to the emerging "Internet of Things (IoT)." Since such embedded devices are tightly coupled both with their environments and with each other through their wireless communication channels, both variations in their environments and the system's need to respond (sometimes rapidly) to those variations may produce (1) the need for such devices to communicate and (2) with it the potential for channel contention to arise, dynamically at run-time. Thus, how wireless channels among the embedded devices are allocated and managed in these systems may significantly influence both communication-specific quality-of-service (QoS) properties (such as message throughput) and broader QoS properties (such as timeliness of system responsiveness) that depend on them.**

A growing body of research has focused on managing different aspects of wireless communication, but has done so mainly in an ad hoc manner, with respect to individual aspects rather than multiple aspects and their potential interactions. Even less attention has been paid to formal methods for assessing how combinations of aspects may influence communication performance, and how to characterize, adapt to, and exploit their combined effects, which is essential to address the challenges noted above.

To overcome these limitations of the current state of the art, this paper makes three main contributions to wireless communication for distributed embedded systems with QoS constraints. First, it shows how a basic but fundamental set of channel admission and modulation decisions can be combined within a single Markov decision process (MDP) model to optimize (in expectation) objectives such as message throughput, even with stochastic arrival and interference characteristics. Second, it identifies regular structure in the value-optimal policies generated off-line from these models, which forms the basis for efficient and accurate heuristics suitable for on-line use. Third, it shows how single- and multi-variable regression techniques can be used to characterize key parameters that govern such regular structure, which then are used to instantiate those heuristics.

I. Introduction

Many new classes of distributed embedded systems are substantially composed of embedded devices that need to communicate with each other, as well as with traditional servers and other system infrastructure. *Wireless* communication among these devices is often also essential, due to deployment costs and constraints.

Wireless communication technologies and standards are increasingly being used to federate embedded devices in a variety of distributed systems application domains, ranging from wind and water monitoring [9], to solar power generation forecasting [2], to industrial process control [27]. In such systems, embedded devices are both tightly coupled with their environments and with each other.

Environmental factors, including background electromagnetic noise at various radio frequencies, or variations in a physical process the system is monitoring or controlling, may dynamically increase or decrease the need for embedded devices to communicate. Thus, when wireless communication is needed, a system may experience both a greater demand for throughput and greater contention for the available wireless communication channels, at once.

Therefore, how wireless channels shared among embedded devices are allocated and managed, in turn can have a significant influence on system behavior. A growing body of research, which we describe further in Section II, has focused on managing individual aspects of wireless communication. However, less attention has been paid to formal approaches aimed at examining and understanding (1) how managing different combinations of aspects may affect communication performance; (2) how those aspects may interact; and (3) how to adapt to, and exploit the combined effects of multiple managed aspects at once, on-line.

Consider the following example scenario in which the system's operating environment is enclosed and under the control of a single "owner" (e.g., an operating room, a smart home, or a factory floor). As a result, the owner can impose operating constraints on, or even actively manage, the radio frequency (RF) spectrum ranges that devices use within the controlled environment. This kind of scenario becomes even more relevant when one considers the fact that "cognitive radio" technology now allows unprecedented control of individual wireless transmissions [18], [22]. A host of specific details (e.g., traditional ones like frequency and power, as well as newly enabled ones like modulation techniques and waveforms) can now be decided on a case by case basis, even for each individual transmission.

The fine-grained control over individual wireless transmissions that is enabled by cognitive radio technology, and the realistic opportunity to exclude transmitters that do not "play by the rules," together imply the question, "what should those rules be?" That is, how should an owner of a portion of radio frequency spectrum manage the use of that spectrum for maximum benefit in terms of different system QoS properties?

This paper takes an initial foray into addressing that question, by using formal optimization theory. Specifically, we model an example spectrum management problem using Markov decision process theory, from which we generate value-optimal [33] decisions (in this paper, those that in expectation optimize throughput) for a pair of aspects: admission and choice of modulation technique.

The rest of this paper is structured as follows. Section II describes related work on managing wireless communication in embedded systems, and on using MDP models and the policies they generate to manage embedded system resources. Section III then examines how channel admission and modulation decisions can be combined within a single Markov decision process (MDP) model through which off-line policy generation can optimize (in expectation) system QoS objectives (specifically, message throughput), even with stochastic arrival and interference characteristics. Section IV identifies and describes regular structure within the policies generated by these models. Section V then shows that a common abstract heuristic algorithm for efficient and accurate on-line decision making can be obtained, which then can be parameterized through a detailed characterization of the specific structure of the policies generated from a given MDP, for on-line use. Finally, Section VI presents conclusions and outlines future directions for this research.

II. Related Work

Wireless communication has received increasing attention within the embedded systems community over the past decade, and wireless distributed embedded systems are now being deployed in a variety of application domains. Du et al. [9] used wind monitors to study hydrodynamics and water quality in urban reservoirs. Achleitner et al. [2] used solar irradiance sensors for power generation forecasting. Sha et al. [27] developed an embedded systems testbed and conducted comparative studies of network protocols based on the WirelessHART standard for industrial process control, which showed that graph routing was able to optimize latency and energy efficiency better than source routing.

Multiple QoS properties such as timeliness, energy constraints, scalability, and resilience are important considerations for wireless networked embedded systems, and a variety of approaches to address trade-offs for different wireless communication scenarios have been developed. Lu et al. developed RAP [19], a wireless communication approach based on the idea of a *velocity monotonic* scheduling policy that considers both timing and distance constraints. The SPEED protocol developed by He et al. [15] uses a combination of feedback control and non-deterministic geographic forwarding to maintain desired message delivery speeds throughout a distributed wireless network. Santinelli et al. examined how computational tasks and message transmissions can be co-scheduled to reduce overall energy consumption [25]. Research by Mazumder and Hallstrom produced an approach for efficient reprogramming of wireless sensor networks [20] that can achieve significant reductions in energy costs when system code must be updated. Chen and Chou addressed the potential of disconnection, interference, or other disruptive events to suspend communication indefinitely, while still ensuring resumption of a propagating update with limited redundant communication [6].

Admission control has seen decades of work in the RF domain [3], [5], [10], [11], [30], including work combining admission and power control [16]. While the RF channel models that guide such control mechanisms are often quite sophisticated, the algorithms themselves are frequently ad hoc, without a basis in formal modeling. In the research described in this paper, we investigate the use of formal Markov decision theory to manage both admission and modulation decisions.

Markov decision processes have been used extensively to automate system management decisions [1], [28], particularly those that involve stochastic elements [4]. Djonin et al. [8] and Zhao et al. [34] have used MDPs to guide secondary use of wireless spectrum (i.e., to minimize interference for primary users). In the context of wireless embedded systems, Lung et al. [7] have used MDP models to optimize battery life in mobile phones. In our previous work [21] we used MDP models to determine wireless communication channel admission decisions, and compared performance predictions from the MDP models with those from a discrete-event simulation model.

Our work most closely follows the approaches taken by Glaubius et al. [12], [13] and Tidwell et al. [31], [33], which applied MDP models to inform resource scheduling decisions in real-time embedded systems where resource allocation durations are stochastic. Our research follows a similar path, but focuses on wireless communication semantics rather than non-preemptive scheduling semantics, which results both in different features in our MDP model and different challenges in defining and applying it. A key distinction in our approach is that the number of transmitters in a wireless embedded system is often limited, which results in a state space that is inherently finite. Although the initial work by Tidwell et al. defined an artificial time horizon that restricted the size of their state space [32], that restriction resulted in modeling errors and policy artifacts that were removed in their later work by folding the inherently infinite state space to produce a finite but exact model. In contrast, in this research the boundedness of the state space is an inherent feature, and the policy artifacts resulting from it, which we describe in more detail in Section IV, are a natural consequence thereof.

III. System Model

In this section we first describe our model of the physical communication process, in which a fixed number of transmitters send messages over wireless (radio) channels. We then define a Markov decision process (MDP) model through which to generate policies channel admission and modulation decisions.

A. Physical Model

In our physical model, the wireless medium is a range of radio frequencies divided into some fixed number of channels, C, and a centralized manager that makes decisions about the use of those channels. The channels are rigid, non-overlapping regions of the wireless (radio frequency) spectrum for which a centralized manager makes both admission and modulation decisions. Although extending our approach to decentralized decision making is a potential direction for future work, in this paper we focus solely on a centralized model because channel contention is most likely to occur among nearby transmitters.

Messages arrive from transmitters via a Poisson process (utilizing a separate *control channel*), are allocated to a channel if admitted, and depart the system if not admitted (i.e., we do not model retries). The manager is responsible for making admission decisions, which are delivered to the appropriate transmitters (again via a separate control channel that is not explicitly modeled). In addition, for each message, the centralized manager selects among two modulation types, which are described below.

We denote the mean message arrival rate by λ, and message durations are assumed to be exponentially distributed with mean $1/\mu$ (following the convention in queuing theory that μ is a *service rate*). The total rate of departure at any specific time, therefore, is proportional to the number of messages in the system (i.e., it is a traditional birth-death process of the type described by Kleinrock [17]). Note that the system performance will be insensitive to the distributional assumptions we make for message duration, since it is an Erlang-loss system [26].

We assume that sufficient power control management is in place so that the signal power for each transmitter (for either modulation technique) is the same at the receiver (i.e., transmitters at a farther distance have a greater transmit power). The noise power is a combination of background (environmental) noise and the interference noise generated by other transmitters that are using common regions of the wireless spectrum.

B. MDP Model

In this work we adopt the definition used by Glaubius et al. [13] of a (discrete-time) Markov decision process as a 5-tuple $(\mathcal{X}, \mathcal{A}, T, R, \gamma)$, with *states* designated as $\chi \in \mathcal{X}$, *actions* designated as $a \in \mathcal{A}$, and a transition system, T, which gives the probability $P_T(\chi' \mid \chi, a)$ of transitioning from state χ to state χ' on action a. The reward function $R(\chi, a, \chi') \in \mathbb{R}_{\geq 0}$ describes the reward that accrues when transitioning from state χ to state χ' via action a, under a discount factor, γ, to ensure convergence of the long term reward.

We use such MDP models to guide both admission and modulation decisions. The two types of modulation we consider here are Orthogonal Frequency Division Multiplexing (OFDM), in which each message consumes only a single channel, and Spread Spectrum (SS), in which each message's transmission is spread across the entire region of wireless spectrum under consideration. The number of OFDM transmitters that can be concurrently transmitting is limited by the number of channels in the wireless spectrum, C. Similarly, the number of SS transmitters that can be concurrently transmitting is limited by the number of pseudo-random codes provided by the SS modulation code set. Without loss of generality, in this work we assume that this limit is also C. In a given region of wireless spectrum, and assuming state-of-the-art modulator design for both OFDM and SS, the size of the code set for SS will likely be close to the number of channels for OFDM in any event [23]. The above two limits effectively bound the total number of concurrent transmissions (to their sum) and naturally bound the extent of the MDP's state space as described below. Extending our approach to additional modulation techniques appears straightforward based on the results we have obtained with OFDM and SS modulation: dedicating another dimension of the state space to each additional modulation technique, as we do already for OFDM and SS; however, doing so is deferred to future work.

1) Transition System: The state space transition diagram for a simple illustrative two-channel four-transmitter MDP is shown in Figure 1. States $(y_2, y_1) \in \mathcal{X}$ represent the number of transmitters using each modulation type. Here, y_2 (shown vertically in the figure) is the number of SS transmitters, constrained to be in the range $0 \leq y_2 \leq C$, and y_1 (shown horizontally) is the number of OFDM transmitters, constrained to be in the range $0 \leq y_1 \leq C$. The total number of transmitters currently in the system (i.e., transmitting) is therefore their sum, $y_2 + y_1$.

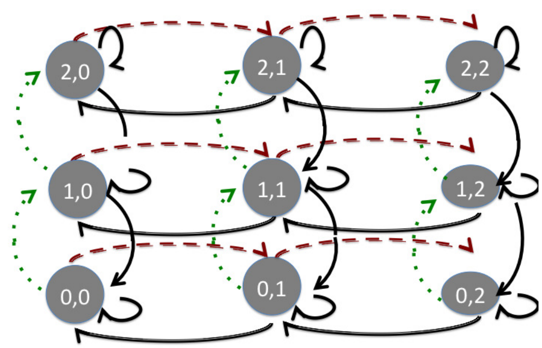

Fig. 1. State space diagram for a simple two-channel four-transmitter MDP. The number of SS transmitters is shown vertically and the number of OFDM transmitters is shown horizontally. Self-loops represent '`no accept`' actions, vertical rising edges represent '`accept SS`' actions, horizontal edges to the right represent '`accept OFDM`' actions, and vertical falling edges and horizontal edges to the left represent message departures.

The MDP is designed to make both admission and modulation decisions. To support this, the available actions are '`accept SS`', '`accept OFDM`', and '`no accept`'; thus $\mathcal{A} = \{\texttt{accept SS}, \texttt{accept OFDM}, \texttt{no accept}\}$.

Upon a message arrival, if the action is to '`accept SS`' then the system transitions one state upward (note that this action is not available when $y_2 = C$). If the action is to '`accept OFDM`' then the system transitions one state to the right (again, this action is not available when $y_1 = C$). If the action is '`no accept`', then the system stays in the same state (follows a self-loop).

Upon a departure, the system transitions down or to the left, depending on the type of transmitter involved (SS vs.

OFDM respectively) but independent of action (formally, all three actions exercise the same transition). The rates are the traditional birth-death rates [17]: λ for arrivals, $y_2 \cdot \mu$ for SS departures and $y_1 \cdot \mu$ for OFDM departures. This continuous-time model is converted to a discrete-time MDP using the uniformization techniques described by Grassmann [14] with uniform rate parameter δ which is set to be greater than the largest rate in the continuous-time model.[1]

2) Reward Function: The reward function, R, characterizes the throughput achievable on an individual channel via the classic Shannon capacity theorem [29], which provides a measure of the channel capacity as a function of the available bandwidth and the signal-to-noise ratio

$$C_s = \gamma_e B_C \log_2(1 + S/N), \tag{1}$$

where C_s is the achievable channel capacity (in bits/s), γ_e is the modulation efficiency, B_C is the channel bandwidth, S is the average signal power (at the receiver), and N is the average noise power. When multiple modulation types are present, we assume that the "signal" power for an alternate message overlapping in the same spectrum is perceived by the receiver as noise when processing the message of interest.

As was noted in Section III-A, each transmitter has the same signal power, and interference occurs only due to background noise and interference from other transmitters. This means that (1) OFDM transmitters do not share channels, (2) SS transmitters all use distinct codes so that they appear as pseudo-random noise to each other and to OFDM transmitters, (3) the interference noise power per transmitter is effectively power-controlled in the same way as the intended signal, and (4) noise power is uniformly distributed across each channel.

For each admitted transmitter, Equation (1) provides the capacity for that individual transmitter. The reward is computed as the number of transmitters allocated to transmit using OFDM modulation multiplied by the associated capacity within each channel plus the number of spread spectrum transmissions multiplied by their individual capacity. This reward, R, can be expressed as follows,

$$R = y_2 \cdot C_{SS} + y_1 \cdot C_{OFDM} \tag{2}$$

where $C_{SS} = C_s/\delta$ for the spread spectrum transmitters, $C_{OFDM} = C_s/\delta$ for the OFDM transmitters, and δ is the uniform rate parameter, which comes out of transforming the continuous-time Markov model into a discrete-time Markov model [14] and therefore changes the units of channel capacity, C_s, from bits/s to simply bits (i.e., δ has units s^{-1}).

3) Illustrative Example: We now illustrate this reward function using the two-channel four-transmitter model shown in Figure 1, where each transmitter is capable of either OFDM or SS modulation. Here, the signal is assumed to fill the spectrum perfectly (i.e., modulation efficiency, γ_e, is assumed to be unity). There are nine total states; we examine each state in turn and formulate an expression for the reward by first articulating values for C_{SS} and C_{OFDM} in terms of the received signal from the transmitter of interest, S, the environmental noise, N, and the signal due to any interfering transmitters. When interfering transmitters are included in the expression, the "noise" due to these transmitters is accounted for by additional instances of S.

The states are organized as in Figure 1, with the horizontal dimension, y_1, representing the OFDM modulated signals and the vertical dimension, y_2, representing the SS modulated signals. Starting with state $(0,0)$, we progress horizontally across the bottom row examining first the OFDM modulated channels (no SS modulation), then vertically up the left-most column allocating SS modulated channels only (no OFDM modulation), and then we examine the remaining states.

State $(0,0)$ indicates that all transmitters are off and therefore the reward is zero. In this case, both $C_{SS} = 0$ and $C_{OFDM} = 0$.

State $(0,1)$ represents the circumstance where one OFDM transmitter is on: the reward is given by $C_{SS} = 0$ and $C_{OFDM} = B_C \log_2((S/N) + 1)$. Understanding that there is only one transmitter enabled using OFDM modulation (i.e., one transmitter's worth of signal, S, only environmental noise, N, and no interfering transmitters), the overall reward is therefore

$$R = 1 \cdot C_{OFDM} = B_C \log_2((S/N) + 1). \tag{3}$$

State $(0,2)$ represents the case where two OFDM transmitters are on. The reward is given by $C_{SS} = 0$ and $C_{OFDM} = B_C \log_2((S/N)+1)$. The capacity of each channel is the same as above since they are independent. Since there are two channels in use, the overall reward is therefore

$$R = 2 \cdot C_{OFDM} = 2B_C \log_2((S/N) + 1). \tag{4}$$

Moving vertically up the left-most column, state $(1,0)$ indicates that one SS transmitter is sending signal power S spread over two channels ($2B_C$). The capacity of the spectrum to transmit the information is given by $C_{SS} = 2B_C \log_2((S/2N) + 1)$, with the environmental noise of 2 channels represented by $2N$. Since there are no OFDM transmitters, $C_{OFDM} = 0$ and overall reward is therefore

$$R = 1 \cdot C_{SS} = 2B_C \log_2((S/2N) + 1). \tag{5}$$

State $(2,0)$ represents a pair of SS transmitters, each spread over the 2 channels of spectrum. Here, the OFDM transmission experiences both the environmental noise in its assigned channel, N, and one half of the signal power of the spread spectrum transmitter, $0.5S$; this gives $C_{OFDM} = B_C \log_2\left(\frac{S}{N+0.5S} + 1\right)$. Similarly, the SS transmission experiences both environmental noise (across both channels in this case, $2N$) and all of the signal power of the OFDM transmitter, S. This yields $C_{SS} = 2B_C \log_2\left(\frac{S}{2N+S} + 1\right)$. The overall

[1] In the literature the uniform rate parameter is frequently represented by γ, but we reserve that symbol for use in the MDP definition.

reward function is therefore

$$\begin{aligned} R &= 1 \cdot C_{SS} + 1 \cdot C_{OFDM} \quad (6) \\ &= 2B_C \log_2\left(\frac{S}{2N+S}+1\right) + \\ &\quad B_C \log_2\left(\frac{S}{N+0.5S}+1\right). \end{aligned}$$

Each state is distinguished by the expression for the denominator in the Shannon capacity (i.e., the noise experienced by each transmission). In Table I we summarize the last three states (shown in the upper-right corner of Figure 1), giving the expressions for C_{SS} and C_{OFDM} for each case.

TABLE I
CAPACITIES FOR THE UPPER-RIGHT STATES.

State	C_{SS}	C_{OFDM}
$(2,1)$	$2B_C \log_2\left(\frac{S}{2N+2S}+1\right)$	$B_C \log_2\left(\frac{S}{N+S}+1\right)$
$(1,2)$	$2B_C \log_2\left(\frac{S}{2N+2S}+1\right)$	$B_C \log_2\left(\frac{S}{N+S}+1\right)$
$(2,2)$	$2B_C \log_2\left(\frac{S}{2N+3S}+1\right)$	$B_C \log_2\left(\frac{S}{N+S}+1\right)$

C. Policy Generation

A policy π for an MDP $(\mathcal{X}, \mathcal{A}, T, R, \gamma)$ maps states in $\mathcal{X}$ to actions in $\mathcal{A}$. Policies are generated by repeated exploration of the MDP's transition system, during which at each discrete decision epoch k the state of the MDP χ_k, is observed and an action $a_k = \pi(\chi_k)$ is chosen. The MDP then transitions to state χ_{k+1} with probability $P_T(\chi_{k+1} \mid \chi_k, a_k)$ and yields immediate reward[2] $r_k = R(\chi_k, a_k, \chi_{k+1})$.

The value of a policy, denoted by V^π, is the expected sum of long-term discounted rewards obtained while following that policy, under discount factor $\gamma \in [0, 1)$, which in these studies we chose to be 0.99 as in Tidwell et al. [33].

$$V^\pi(\chi) = E\left\{\sum_{k=0}^{\infty} \gamma^k r_k \mid \chi_0 = \chi, a_k = \pi(\chi_k)\right\} \quad (7)$$

We also adopt the terminology from Tidwell et al., denoting a policy as being *value-optimal* [33] (and thus it may be selected as the output of the policy generation process) if it maximizes long term value in expectation when computed for an MDP with a finite number of states and actions [24].

IV. Parametric Study of Generated Value-Optimal Policies

To validate the policies generated by the MDP-based approach described in Section III, and to assess which features of our system model have the greatest impact on the resulting policies, we performed a parametric study of value-optimal policies generated by the MDP, varying each of the following parameters individually: offered load (defined as λ/μ), ranging from 0.2 E (erlangs) to 2.4 E; modulation efficiency of each modulation type, ranging from 0.7 to 1.0; and the signal-to-noise ratio (SNR) within the wireless spectrum, ranging from 1.0 to 12.0. The study was performed on a rectangular state space with 16 OFDM channels and 16 SS codes. Figure 2 shows an example of a generated value-optimal policy, for the following parameterization: offered load of 0.6 E, both modulation efficiencies at 1.0, and signal-to-noise ratio of 2.0.

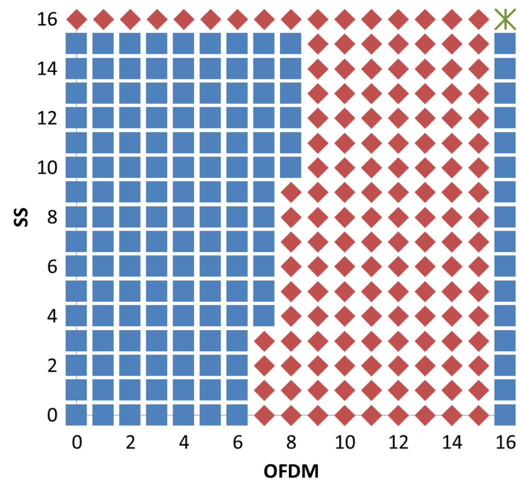

Fig. 2. Value-optimal policy for numbers of SS and OFDM transmitters ranging from 0 to 16, offered load = 0.6, SS efficiency = 1.0, OFDM efficiency = 1.0, and SNR = 2. Squares represent the '`accept SS`' action, diamonds the '`accept OFDM`' action, and the asterisk represents the '`no accept`' action.

A. Structural Observations

A number of observations can be made about this policy that are primarily functions of the structure of the state space. As a result, they are also relevant to the other policies generated in this study. First, in the upper right-hand corner, the action is '`no accept`' because there is no other action available. Second, along the top row the MDP can only choose between '`accept OFDM`' and '`no accept`' actions because the '`accept SS`' action is unavailable. In this circumstance, the value-optimal policy is to choose the '`accept OFDM`' action. Third, along the right-most column, the MDP can only choose between '`accept SS`' and '`no accept`' actions because the '`accept OFDM`' action is unavailable. We call the area of the MDP state space that is identified by these three observations the *restricted region*, since the actions available are limited.[3]

Fourth, the remaining portion of the state space is called the *unrestricted region*, since all actions are available to the MDP. Here, when the number of OFDM transmitters is below a certain level, the value-optimal action is to accept more SS transmitters, and above this level the value-optimal action is to accept more OFDM transmitters. We see a similar pattern consistently across the parameter space that was explored in this study.

[2] Note that in the MDP described above, the reward is only a function of the destination state, $r_k = R(\chi_{k+1})$.

[3] As we noted in Section II, this is a natural consequence of the bounded number of transmitters supported by our system model.

We next present the results of the full parametric study, with individual cases illustrated in Figures 2 through 9.

B. Offered Load

Figure 3 shows the effect of an increased offered load (moving from 0.6 E in Figure 2 to 1.8 E in Figure 3). Modulation efficiencies have been kept at 1.0 and SNR is still at 2.0.

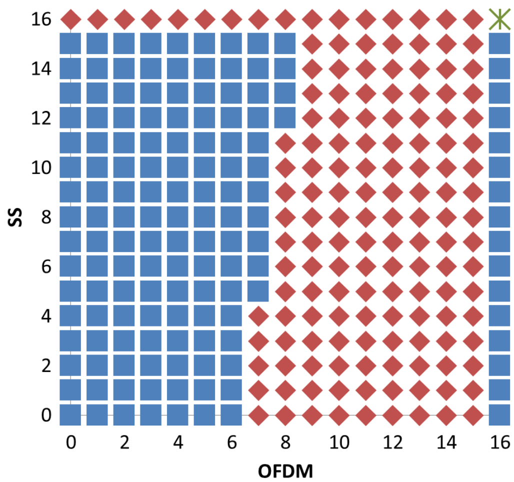

Fig. 3. Value-optimal policy for numbers of SS and OFDM transmitters ranging from 0 to 16, offered load = 1.8, SS efficiency = 1.0, OFDM efficiency = 1.0, and SNR = 2. Squares represent the '`accept SS`' action, diamonds the '`accept OFDM`' action, and the asterisk represents the '`no accept`' action.

We can see in the figure that, generally, offered load has a limited impact. The value-optimal actions in the restricted regions have not changed. The slope of the boundary line in the unrestricted region has changed a bit, but only slightly. While the impact is small, it is still meaningful, and so the effect of offered load is included in the boundary line determination techniques described in Section V.

C. Modulation Efficiency

Returning to the baseline configuration of Figure 2, we investigate the impact of modulation efficiency by first reducing SS efficiency to 0.9 in Figure 4 and then to 0.8 in Figure 5.

While the value-optimal actions in the restricted regions have not changed, the impact of modulation efficiency in the unrestricted region is dramatic. As the spread spectrum modulation efficiency decreases, the boundary between the '`accept SS`' action and the '`accept OFDM`' action moves to the left, indicating that the value-optimal choice is to use more OFDM modulation and less SS modulation.

In a similar manner, we investigate the impact of OFDM modulation efficiency by reducing OFDM efficiency from the 1.0 of Figure 2 to 0.9 in Figure 6 and then to 0.8 in Figure 7.

Once again, the value-optimal actions in the restricted regions have not changed, but the impact of modulation efficiency in the unrestricted region is dramatic. As the OFDM modulation efficiency decreases, the boundary between the

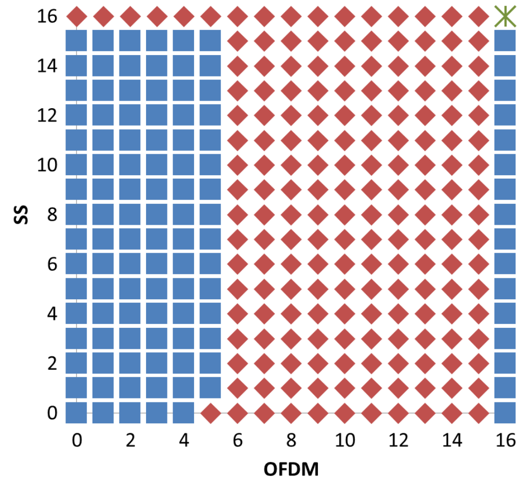

Fig. 4. Value-optimal policy for numbers of SS and OFDM transmitters ranging from 0 to 16, offered load = 0.6, SS efficiency = 0.9, OFDM efficiency = 1.0, and SNR = 2. Squares represent the '`accept SS`' action, diamonds the '`accept OFDM`' action, and the asterisk represents the '`no accept`' action.

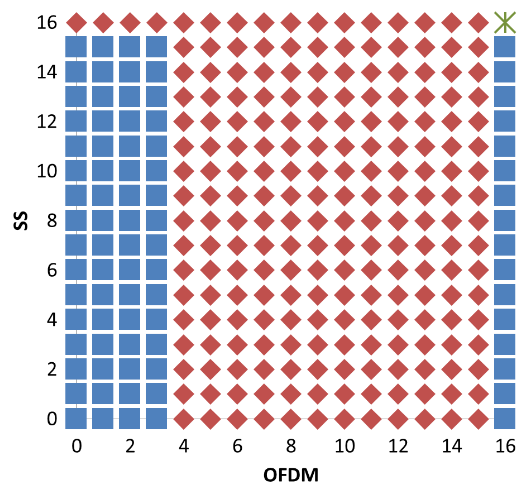

Fig. 5. Value-optimal policy for numbers of SS and OFDM transmitters ranging from 0 to 16, offered load = 0.6, SS efficiency = 0.8, OFDM efficiency = 1.0, and SNR = 2. Squares represent the '`accept SS`' action, diamonds the '`accept OFDM`' action, and the asterisk represents the '`no accept`' action.

'`accept SS`' action and the '`accept OFDM`' action moves to the right, indicating that the value-optimal choice is to use more SS modulation and less OFDM modulation. Interestingly, the slope of the boundary changes more dramatically in this case.

D. Signal-to-Noise Ratio

Finally, we investigate the impact of signal-to-noise ratio by comparing Figure 2, which has an SNR of 2.0, with Figures 8 and 9, which increase the SNR, first to 4.0 and then to 8.0. In all three of these figures, the offered load is 0.6 E and the modulation efficiencies are set to 1.0.

Here, for the first time, we see changes in the value-optimal

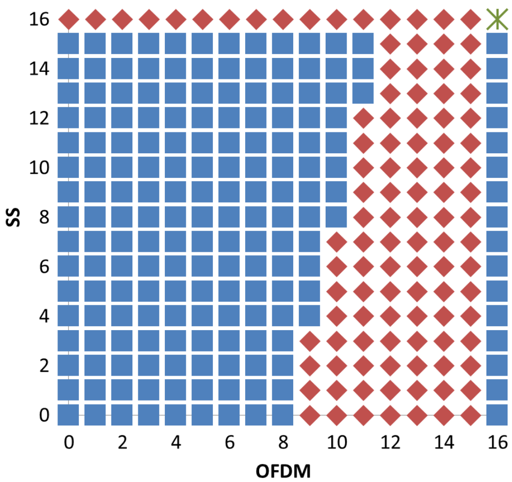

Fig. 6. Value-optimal policy for numbers of SS and OFDM transmitters ranging from 0 to 16, offered load = 0.6, SS efficiency = 1.0, OFDM efficiency = 0.9, and SNR = 2. Squares represent the '`accept SS`' action, diamonds the '`accept OFDM`' action, and the asterisk represents the '`no accept`' action.

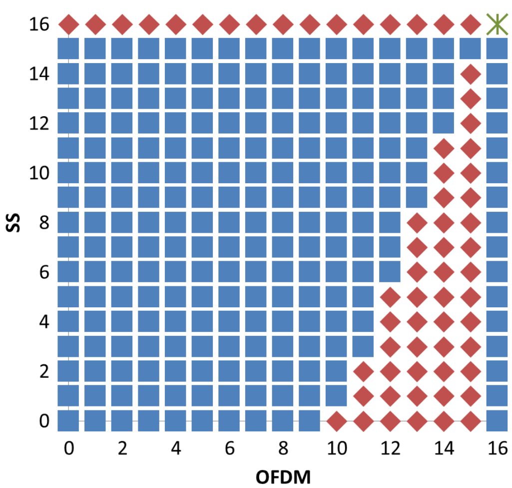

Fig. 7. Value-optimal policy for numbers of SS and OFDM transmitters ranging from 0 to 16, offered load = 0.6, SS efficiency = 1.0, OFDM efficiency = 0.8, and SNR = 2. Squares represent the '`accept SS`' action, diamonds the '`accept OFDM`' action, and the asterisk represents the '`no accept`' action.

actions in the restricted regions. As the SNR increases, in an increasing number of states the long term reward is best served by not admitting an additional transmitter. This is clearly a circumstance in which the long term reward is better served by a different action than the "greedy" approach of seeking immediate reward. As the SNR varies, the impact in the unrestricted region is also interesting. While there aren't dramatic shifts in the number of states for which the value-optimal action is to transmit with one modulation type vs. the other, the slope of the line that divides the region is noticeably impacted by SNR.

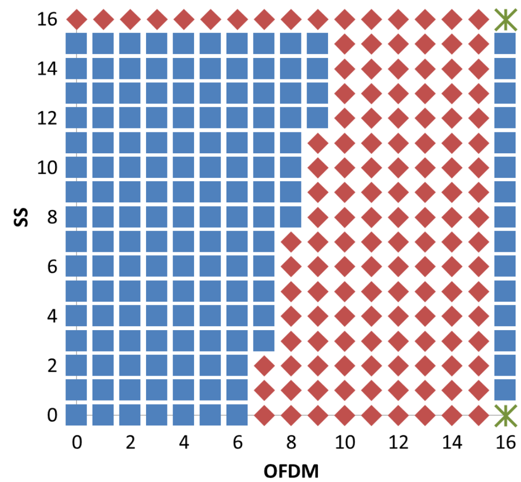

Fig. 8. Value-optimal policy for numbers of SS and OFDM transmitters ranging from 0 to 16, offered load = 0.6, SS efficiency = 1.0, OFDM efficiency = 1.0, and SNR = 4. Squares represent the '`accept SS`' action, diamonds the '`accept OFDM`' action, and the asterisks represent the '`no accept`' action.

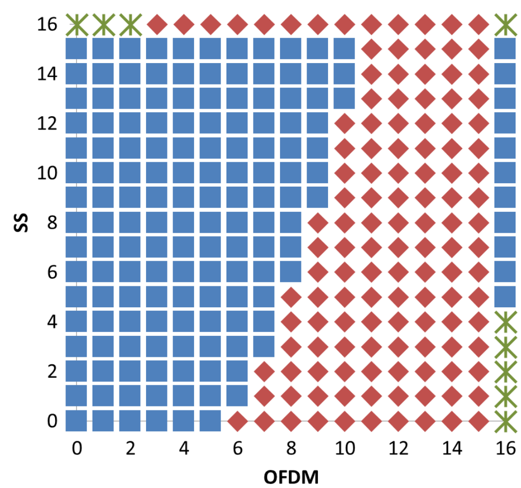

Fig. 9. Value-optimal policy for numbers of SS and OFDM transmitters ranging from 0 to 16, offered load = 0.6, SS efficiency = 1.0, OFDM efficiency = 1.0, and SNR = 8. Squares represent the '`accept SS`' action, diamonds the '`accept OFDM`' action, and the asterisks represent the '`no accept`' action.

E. Summary

We now summarize the results of the full parametric study. We note that each value-optimal policy has distinct patterns in the restricted region of the state space vs. the unrestricted region, and that in each region there are regularly structured boundaries between groups of states in which the same value-optimal action is taken. In Section V we show how these boundaries between action choices in each region of the state space, can be exploited to develop efficient on-line heuristics for admission and modulation decisions.

In the unrestricted region, the value-optimal actions are effectively divided by a linear boundary and all actions to the left of the boundary are ‘`accept SS`’ and all actions to the right of it are ‘`accept OFDM`’. This line can be characterized in terms of the set of parameters described above, as we discuss further in Section V.

Modulation efficiency has a substantial impact on both the position and slope of this linear boundary. It is not surprising that the boundary moves in response to changes in modulation efficiency, though as we note in Section V, it is important to characterize quantitatively how the boundary moves.

The SNR primarily impacts the slope of the boundary line. This again points to the need for a quantitative understanding of how the linear boundary is positioned as a function of the input parameters.

In the restricted region, the signal-to-noise ratio has a significant effect, establishing a boundary between ‘`accept`’ and ‘`no accept`’ actions. Starting at an SNR of 4.0, the best (value-optimal) action is not always to accept a new message upon arrival, and as the SNR increases, the number of states with value-optimal action ‘`no accept`’ increases as well. Which ‘`accept`’ action is allowed is restricted by the dimension of the restricted region that is involved (either top row or right-most column).

V. Heuristics

In Section III, we described the use of MDP models to generate value-optimal policies for managing wireless spectrum, specifically guiding admission and modulation decisions. However, the exponential cost of computing a value-optimal policy is likely to be impractical in an on-line setting. For example, as Figure 10 illustrates, on a 2.3 GHz Opteron machine with 16 GB of memory, while policy generation for a 4 channel model took less than a second, a 32 channel model required hours, and larger models would need days or even weeks, which clearly is unsuitable for on-line use.

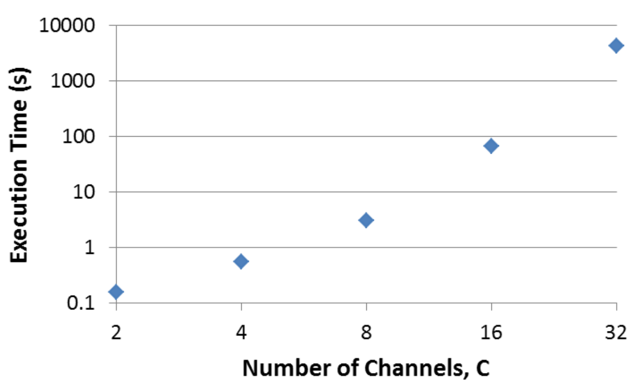

Fig. 10. Execution time required to compute value-optimal policies vs. state space size.

To address this concern, we now develop simple and efficient but effective heuristics that approximate the value-optimal policies of the MDP models we consider. These heuristics are based on the discovery of efficiently computable boundaries between regions that are characterized by a common action, which we described in Section IV. Specifically, in our investigation of the parameter space, each and every value-optimal policy has the following properties: (1) the action in the upper-right corner is to not accept (the only action available); (2) in the top-most row the action in states of the left of a boundary is ‘`no accept`’ and in states to the right of it is ‘`accept OFDM`’; (3) in the right-most column the action for states below a boundary is ‘`no accept`’ and for states above that boundary is ‘`accept SS`’; and (4) actions for other states (in the unrestricted region) are divided by a line that separates SS from OFDM modulation.

Note that the various parameters impact the position and orientation of the boundaries, but do not alter their form. Based on these properties, we define a heuristic approach to approximating the value-optimal policies for the parameter space as follows. First, using off-line analysis, we describe the boundaries as a function of the parameters. For the unrestricted region, given the offered load, OL, modulation efficiency, $\gamma_{e.SS}$ and $\gamma_{e.OFDM}$, and signal-to-noise ratio, r_{SN},

$$B_U(y_2, y_1 \mid OL, \gamma_{e.SS}, \gamma_{e.OFDM}, r_{SN}) = 0 \qquad (8)$$

represents the boundary between where the value-optimal decision is ‘`accept SS`’ vs. ‘`accept OFDM`’. Here, y_2 is the number of SS transmitters and y_1 is the number of OFDM transmitters. For the restricted region, given the signal-to-noise ratio, r_{SN}, $B_{RS}(r_{SN})$ and $B_{RF}(r_{SN})$ are boundaries between where where the value-optimal decision is ‘`accept SS`’ vs. ‘`no accept`’ for B_{RS} and ‘`accept OFDM`’ vs. ‘`no accept`’ for B_{RF}.

Second, the on-line algorithm compares the current state to the appropriate boundary, and determines the action to take based on to which side of the boundary it finds itself. We first describe the on-line algorithm, given the parameterized boundaries that separate the state space into regions of distinct action. We then describe how we can calibrate the boundaries.

A. Towards an On-line Heuristic Algorithm

We can approximate the value-optimal policy via the on-line algorithm shown in Figure 11. Here, the boundary line of Equation (8), $B_U(y_2, y_1) = 0$, is expressed in the form

$$y_1 = m \cdot y_2 + b \qquad (9)$$

where m and b are coefficients that are determined using the techniques described below. Note that this boundary formulation expresses the horizontal component of the state as a function of the vertical component (i.e., the opposite of what is traditional). This helps us with vertical lines (which we have) and would be problematic with horizontal lines (which we don’t have). It is equivalent to transposing the earlier figures and showing the number of spread spectrum transmitters, y_2, horizontally and the number of OFDM transmitters, y_1, vertically.

In the algorithm, lines (4) and (10) each compare the current state with the appropriate boundary in the restricted region,

```
(1)  if y1 = C and y2 = C then
(2)       a ← 'no accept'
(3)  else if y2 = C then
(4)       if y1 ≤ B_RF then
(5)            a ← 'no accept'
(6)       else
(7)            a ← 'accept OFDM'
(8)       endif
(9)  else if y1 = C then
(10)      if y2 ≤ B_RS then
(11)           a ← 'no accept'
(12)      else
(13)           a ← 'accept SS'
(14)      endif
(15) else if y1 ≤ m · y2 + b then
(16)      a ← 'accept SS'
(17) else
(18)      a ← 'accept OFDM'
(19) endif
```

Fig. 11. On-line algorithm to approximate value-optimal policies. C is the number of allowed transmitters of each modulation type, y_2 is the number of SS transmitters, y_1 is the number of OFDM transmitters, and a is the action. Lines (1) to (14) correspond to the restricted region of the state space, and lines (15) to (19) correspond to the unrestricted region.

and line (15) compares the current state with the boundary in the unrestricted region.

1) Determining Unrestricted Region Boundary Line: We express the coefficients of the line B_U in terms of the following parameters: offered load, modulation efficiency, and signal-to-noise ratio. We characterize this relationship via multi-variable linear regression. Using 10 variations of each input parameter within the operational space, we determine the true value-optimal policy for each parameterization. This is used as input into MATLAB's multi-variable regression function `mvregress` (version R2013a), which generates a multivariate general linear model, to determine the coefficients of the line B_U. The ranges of the input parameter values are: offered load, $0.2 \leq OL \leq 2.4$, signal-to-noise ratio, $1 \leq r_{SN} \leq 12$. For the modulation efficiencies, we limited the parameters to the range $0.91 \leq \gamma_e \leq 1$, since modern modulation techniques typically have efficiencies in that range [23]. This results in a 4-predictor regression (offered load, SS efficiency, OFDM efficiency, and signal-to-noise ratio) that yields the coefficients of the line B_U expressed in Equation (9).

2) Determining Restricted Region Boundaries: In the parametric study described in Section IV, we observed that the boundary between accept and no accept decisions in the restricted region appears to be only a function of the signal-to-noise ratio and is not influenced by the other parameters. This hypothesis was examined further by varying more than one parameter at once (the previous studies altered only one parameter at a time). We observed that only SNR influenced the boundaries between actions in the restricted region, and so for that region we only need to perform single variable regression based on the signal-to-noise ratio, r_{SN}.

B. Assessment of the Heuristic Algorithm

We first assess our approximation of the value-optimal policy in the restricted region, and then we assess our approximation of it in the unrestricted region. In the restricted region, we are interested in comparing two thresholds: the boundary between actions in the value-optimal policy, and the threshold used in the on-line algorithm given in Figure 11. Figure 12 compares these two thresholds for the portion of the restricted region in which only OFDM modulation is allowed. The points indicate the value-optimal threshold and the line indicates the heuristic threshold: clearly, the alignment is quite good ($r^2 = 0.97$). For the restricted region in which only SS modulation is allowed, the heuristic is comparably effective ($r^2 = 0.93$).

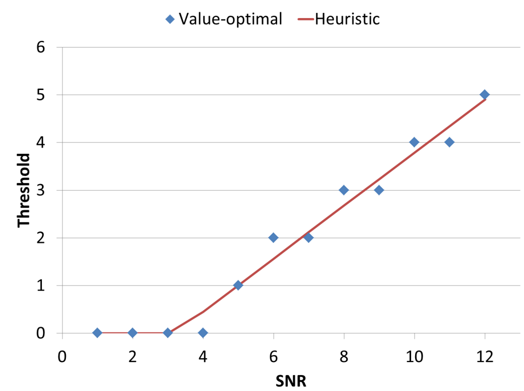

Fig. 12. Value-optimal and heuristic thresholds separating actions in the restricted region in which only OFDM modulation is allowed.

For the unrestricted region, we applied the multi-variable linear regression described above to a training set of 48 configurations (varying each parameter independently). 48 additional randomly generated configurations (constrained to be in the ranges used for training, and not to be in the initial set) were used for testing. In each case, we ran the on-line algorithm given in Figure 11 and compared the chosen action to that in the value-optimal policy generated from the MDP.

Two specific comparisons were made between the heuristic policies and the value-optimal policies. First, we informally assessed the slope and intercept of each policy to ensure that they were similar. Second, we used an r-squared test to measure quantitatively the differences between the heuristic policy and the value-optimal policy. We obtained a value of $r^2 = 0.95$, which indicates close agreement between the value-optimal policy and our heuristic approximation of it.

VI. Conclusions and Future Work

Distributed embedded systems are increasingly dependent on wireless communication among devices. Using message throughput as the metric of interest, we have shown in this paper how a formal Markov decision process (MDP) model can generate value-optimal policies for combined admission and modulation decisions. We also have shown how these value-optimal policies in turn can be approximated using regression techniques to obtain accurate heuristics suitable for use in modern embedded device systems.

Extending our approach to incorporate additional modulation techniques appears straightforward, and considering spread spectrum modulators that do not consume the entire region of wireless spectrum under management appears both worthwhile based on the results we have presented here, and potentially may be needed for the model to remain usable as the number of channels grows. Other promising directions for extending this work include expanding its scope to include additional aspects impacting wireless channels (e.g., fading, multipath interference, distance), which likely would require a more complex channel model to drive the reward function, as would increasing the number of aspects under active management at once (e.g., transmit power for each message, coding choices such as channel codes, error codes, etc.). Extending our approach to decentralized decision making is another important direction for future work, both to increase the applicability of the foundational techniques described here to larger-scale distributed wireless embedded systems, and to explore how the techniques described here may need to be adapted to accommodate larger-scale contexts.

Acknowledgments

This work was supported in part by NSF CNS-1527510.

References

[1] M. Abundo, V. Cardellini, and F. Lo Presti, "An MDP-based admission control for a QoS-aware service-oriented system," in *Proc. of 19th IEEE Int'l Workshop on Quality of Service*, Jun. 2011, pp. 14:1–14:3.

[2] S. Achleitner, T. Liu, A. Kamthe, and A. E. Cerpa, "SIPS: Solar irradiance prediction system," in *Proc. of 13th ACM/IEEE Int'l Conf. on Information Processing in Sensor Networks*, Apr. 2014, pp. 225–236.

[3] I. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "A survey on spectrum management in cognitive radio networks," *IEEE Communications Magazine*, vol. 46, no. 4, pp. 40–48, Apr. 2008.

[4] B. Bethke, L. F. Bertuccelli, and J. P. How, "Experimental demonstration of adaptive MDP-based planning with model uncertainty," in *Proc. of Guidance, Navigation, and Control Conference*, Aug. 2008.

[5] P. Chaporkar and S. Sarkar, "Admission control framework to provide guaranteed delay in error-prone wireless channel," in *Proc. of 58th IEEE Vehicular Technology Conference*, vol. 5, Oct. 2003, pp. 3503–3507.

[6] M.-H. Chen and P. H. Chou, "Telescribe: A scalable, resumable wireless reprogramming approach," in *Proc. of Int'l Conf. on Embedded Software*, Oct. 2010, pp. 139–148.

[7] T. L. Cheung, K. Okamoto, F. Maker, X. Liu, and V. Akella, "Markov decision process (MDP) framework for optimizing software on mobile phones," in *Proc. of Int'l Conf. on Embedded Software*, Oct. 2009, pp. 11–20.

[8] D. V. Djonin, Q. Zhao, and V. Krishnamurthy, "Optimality and complexity of opportunistic spectrum access: A truncated Markov decision process formulation," in *Proc. of IEEE Int'l Conf. on Communications*, Jun. 2007, pp. 5787–5792.

[9] W. Du, Z. Xing, M. Li, B. He, C. H. C. Lloyd, and H. Miao, "Optimal sensor placement and measurement of wind for water quality studies in urban reservoirs," in *Proc. of 13th ACM/IEEE Int'l Conf. on Information Processing in Sensor Networks*, Apr. 2014, pp. 167–178.

[10] H. Fattah and C. Leung, "An overview of scheduling algorithms in wireless multimedia networks," *IEEE Wireless Communications*, vol. 9, no. 5, pp. 76–83, Oct. 2002.

[11] L. Georgiadis, M. J. Neely, and L. Tassiulas, *Resource Allocation and Cross-layer Control in Wireless Networks*. Boston, MA: Now Publishers Inc., 2006.

[12] R. Glaubius, T. Tidwell, B. Sidoti, D. Pilla, J. Meden, C. Gill, and W. Smart, "Scalable scheduling policy design for open soft real-time systems," in *Proc. of IEEE Real-Time and Embedded Technology and Applications Symposium*, Apr. 2010, pp. 237–246.

[13] R. Glaubius, T. Tidwell, W. Smart, and C. Gill, "Scheduling design and verification for open soft real-time systems," in *Proc. of Real-Time Systems Symposium*, Nov. 2008, pp. 505–514.

[14] W. K. Grassmann, "Transient solutions in Markovian queueing systems," *Computers & Operations Research*, vol. 4, no. 1, pp. 47–53, 1977.

[15] T. He, J. Stankovic, C. Lu, and T. Abdelzaher, "SPEED: A stateless protocol for real-time communication in sensor networks," in *Proc. of Int'l Conf. on Distributed Computing Systems*, May 2003, pp. 46–55.

[16] D. Kivanc, G. Li, and H. Liu, "Computationally efficient bandwidth allocation and power control for OFDMA," *IEEE Transactions on Wireless Communications*, vol. 2, no. 6, pp. 1150–1158, Nov. 2003.

[17] L. Kleinrock, *Queueing Systems. Volume 1: Theory*. New York, NY: Wiley-Interscience, 1975.

[18] K. Letaief and W. Zhang, "Cooperative communications for cognitive radio networks," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 878–893, May 2009.

[19] C. Lu, B. Blum, T. Abdelzaher, J. Stankovic, and T. He, "RAP: A real-time communication architecture for large-scale wireless sensor networks," in *Proc. of IEEE Real-Time and Embedded Technology and Applications Symposium*, 2002, pp. 55–66.

[20] B. Mazumder and J. Hallstrom, "An efficient code update solution for wireless sensor network reprogramming," in *Proc. of Int'l Conf. on Embedded Software*, Oct. 2013, pp. 4:1–4:10.

[21] J. Meier, B. Karaus, S. Sistla, T. Tidwell, R. D. Chamberlain, and C. Gill, "Assessing the appropriateness of using Markov decision processes for RF spectrum management," in *Proc. of 16th ACM Int'l Conf. on Modeling, Analysis & Simulation of Wireless and Mobile Systems*, Nov. 2013, pp. 41–48.

[22] J. Mitola and G. Maguire, Jr., "Cognitive radio: Making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[23] J. Proakis, *Digital Communications*. New York, NY: McGraw-Hill, 1995.

[24] M. L. Puterman, *Markov Decision Processes*. Hoboken, NJ: John Wiley & Sons, Inc., 1994.

[25] L. Santinelli, M. Marinoni, F. Prosperi, F. Esposito, G. Franchino, and G. Buttazzo, "Energy-aware packet and task co-scheduling for embedded systems," in *Proc. of Int'l Conf. on Embedded Software*, Oct. 2010, pp. 279–288.

[26] B. A. Sevastyanov, "An ergodic theorem for Markov processes and its application to telephone systems with refusals," *Theory of Probability & Its Applications*, vol. 2, no. 1, pp. 104–112, 1957.

[27] M. Sha, D. Gunatilaka, C. Wu, and C. Lu, "Implementation and experimentation of industrial wireless sensor-actuator network protocols," in *Proc. of European Conference on Wireless Sensor Networks*, Feb. 2015, pp. 234–241.

[28] G. Shani, R. I. Brafman, and D. Heckerman, "An MDP-based recommender system," in *Proc. of 18th Conf. on Uncertainty in Artificial Intelligence*, Aug. 2002, pp. 453–460.

[29] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 1948.

[30] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Transactions on Automatic Control*, vol. 37, no. 12, pp. 1936–1948, Dec. 1992.

[31] T. Tidwell, C. Bass, E. Lasker, M. Wylde, C. Gill, and W. Smart, "Scalable utility aware scheduling heuristics for real-time tasks with stochastic non-preemptive execution intervals," in *Proc. of 23rd Euromicro Conf. on Real-Time Systems*, Jul. 2011, pp. 238–247.

[32] T. Tidwell, R. Glaubius, C. Gill, and W. Smart, "Scheduling for reliable execution in autonomic systems," in *Proc. of 5th Int'l Conf. on Autonomic and Trusted Computing*, Jun. 2008, pp. 149–161.

[33] T. Tidwell, R. Glaubius, C. D. Gill, and W. D. Smart, "Optimizing expected time utility in cyber-physical systems schedulers," in *Proc. of IEEE 31st Real-Time Systems Symposium*, Dec. 2010, pp. 193–201.

[34] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 3, pp. 589–600, Apr. 2007.