Optimal Timer-Based Best Node Selection for Wireless Systems with Unknown Number of Nodes

Rajat Talak, *Student Member, IEEE*, and Neelesh B. Mehta, *Senior Member, IEEE*

***Abstract*—The distributed, low-feedback, timer scheme is used in several wireless systems to select the best node from the available nodes. In it, each node sets a timer as a function of a local preference number called a *metric*, and transmits a packet when its timer expires. The scheme ensures that the timer of the best node, which has the highest metric, expires first. However, it fails to select the best node if another node transmits a packet within Δ s of the transmission by the best node. We derive the optimal metric-to-timer mappings for the practical scenario where the number of nodes is unknown. We consider two cases in which the probability distribution of the number of nodes is either known a priori or is unknown. In the first case, the optimal mapping maximizes the success probability averaged over the probability distribution. In the second case, a robust mapping maximizes the worst case average success probability over all possible probability distributions on the number of nodes. Results reveal that the proposed mappings deliver significant gains compared to the mappings considered in the literature.**

***Index Terms*—Multiple access, selection, timer, opportunistic transmission, collision, robust design.**

I. Introduction

SELECTION is an important technique that is used to enhance the performance of several wireless systems. For example, in a cooperative relaying system, the relay best suited to forward the source's message to the destination is selected [1], [2]. Scheduling, which is used in cellular systems to harness multi-user diversity, is also a form of selection since the user with the highest signal-to-noise ratio (SNR) is selected [3, Chap. 6]. In wireless sensor networks (WSNs), selecting the node that senses is used to improve network lifetime [4], [5]. In vehicular ad hoc networks (VANETs), vehicle selection is used to speed up information dissemination [6]–[8]. Implementing various notions of fairness, such as proportional fairness and max-min fairness, can also be shown to be a selection problem [9].

In all the systems mentioned above, selection occurs as follows. Each node maintains a preference number called a *metric* that is a function of local parameters such as channel gains or measurements of the node. For example, in amplify-and-forward relaying, the metric of a relay is the harmonic mean of the source-to-relay and relay-to-destination channel gains [1], and in proportional fair scheduling, the metric of a node is the ratio of its instantaneous channel gain to its average channel gain [10]. Instead, in [8], the metric is a function of the vehicle's speed and position. The goal of selection is to help a common node called *sink* identify the node with the highest metric, which is called the *best node*.

A fundamental issue with selection is that each node knows only its own metric and not that of the other nodes because the nodes are geographically separated. Therefore, a distributed selection algorithm is needed to identify the best node. The timer-based selection scheme is a popular and important example of a distributed selection scheme [1], [5], [9], [11]. In it, the nodes use a common monotone non-increasing (MNI) metric-to-timer mapping $f(\cdot)$. A node i with metric μ_i sets its timer as $T_i = f(\mu_i)$ and transmits a small *timer packet* when its timer expires. The timer packet typically contains only the node's identity. The MNI property ensures that the first packet to reach the sink is from the best node. The timer scheme is attractive because it is simple to implement and requires limited feedback signaling from the sink.

However, due to its distributed nature, the timer scheme can fail to select the best node. This occurs under the following two scenarios: (i) the timer of the best node does not expire within the stipulated selection duration of $T_{\max}$, or (ii) the timer of the second best node expires within a *vulnerability window* Δ after the expiry of the best node's timer. This results in a collision of the two timer packets, as a result of which the sink fails to decode the timer packet from the best node. Δ is determined by the physical layer capabilities of the system. Typically, it is a sum of the maximum propagation delay, switching time, and maximum time synchronization error. In case the system does not possess carrier sensing capability or if it is susceptible to the hidden nodes problem, Δ also accounts for the timer packet duration [1], [12].

The probability of selecting the best node, which is a fundamental measure of how effective the selection scheme is, depends on the metric-to-mapping. In [11], an optimal timer mapping that maximizes the success probability was shown to be a staircase mapping, in which timers expire only at $\{0, \Delta, 2\Delta, \ldots, N\Delta\}$ or not at all. Here, $N = \left\lfloor \frac{T_{\max}}{\Delta} \right\rfloor$ and $\lfloor\cdot\rfloor$ denotes the floor function. However, [11] assumed that all the nodes know the total number of nodes in the system. However, this is often not the case in practice.

- *Cooperative Decode-and-Forward (DF) Relaying:* Consider a cooperative system in which a source first broadcasts a message to a set of DF relays. One of the relays that decoded the source's message is selected to forward

Manuscript received October 17, 2012; revised April 20 and August 14, 2013. The editor coordinating the review of this paper and approving it for publication was H. Li.

A part of this paper has been presented in the National Conf. on Communications (NCC), Delhi, India, Feb. 2013.

The authors are with the Dept. of Electrical Communication Eng., Indian Institute of Science (IISc), Bangalore (e-mail: rajatalak@gmail.com, nbmehta@ece.iisc.ernet.in).

This research was partly supported by the Broadcom Foundation, USA; Dept. of Science and Technology, India and EPSRC, UK under the auspices of the India-UK Advanced Technology Center (IU-ATC).

Digital Object Identifier 10.1109/TCOMM.2013.091213.120783

the message to the destination. The number of relays that decode the source's message is a function of the source-to-relay channel gains and, hence, is a random variable (RV) that is not known a priori to the source or the destination.

- *Opportunistic Wireless Local Area Networks (WLANs):* Opportunistic media access control (MAC) schemes for WLANs, e.g., [13], exploit multi-user diversity by making nodes with higher SNRs transmit earlier. Since WLAN traffic is bursty, the number of nodes that have packets to send is an RV and is not known to the access point (AP) or any other node in the network.
- *Event Detection in WSNs:* In a WSN, an event may be simultaneously detected by a number of sensor nodes. However, it is sufficient for one node to convey it to the fusion center. However, the fusion center does not a priori know how many nodes detected the event.

The random backoff that is used in [14]–[16] can also be thought of as a timer scheme, albeit a randomized one. However, there are three important differences between these schemes and the selection scheme considered in this paper. In the former, the objective is for the system to receive as many packets as possible from all the nodes in the system and ensure congestion control. On the other hand, in a selection scheme, the objective is to reliably select the best node. Secondly, there is no notion of a best node in [14]–[16]. Thirdly, while we use a deterministic metric-to-timer mapping, the timer is chosen randomly from within a window of values in [14]–[16].

A. Contributions

In this paper, we develop optimal timer mappings when the number of nodes in the system is unknown. This is done for two general models: (i) *unknown number of nodes with a prior distribution that is known to all the nodes*, and (ii) *unknown number of nodes with an unknown distribution*. In the known prior distribution model, we maximize the average success probability, which is the success probability averaged over the prior distribution on the number of nodes. In the unknown distribution model, the only available knowledge is that the number of nodes lies anywhere between $k_{\min}$ and $k_{\max}$, where $k_{\min} \leq k_{\max}$. Here, we maximize the worst case average success probability.

For the known prior distribution model we first show that the optimal timer mapping has a discrete, staircase structure, in which any node's timer expires either at $0, \Delta, 2\Delta, \ldots,$ or not at all. When the number of nodes is a binomial or a Poisson RV, we show that it is characterized by a simple recursion in the number of timer levels. The choice of these distributions, which we shall refer to as *priors*, is motivated by the examples discussed earlier. For example, if all the source-to-relay channel gains are independent and identically distributed (i.i.d.) [1], [2], [17], then the number of DF relay nodes that decode the source's message is a binomial RV. In WSNs, when the sensor nodes are spread uniformly in a geographical area with a density λ, the number of nodes that detect the event can be modeled as a Poisson RV with mean λA, where A is the area across which the event is detected. In WLANs, the number of nodes with a packet to transmit has also been modeled as a Poisson RV in [18].

For the unknown distribution model, we again show that the MNI mapping that maximizes the worst case average success probability, which we shall refer to as the *robust mapping*, has a discrete staircase structure. We develop structural results about the optimal mapping that help compute it easily. Compared to the schemes proposed in the literature, we show that the robust mapping not only achieves the highest worst case success probability, but it also has the narrowest *success probability band*, which is the range of values of the average success probability. This shows its robustness to uncertainty. The optimal and robust mappings also serve as fundamental performance benchmarks for distributed selection schemes.

The timer schemes developed above have applications to emergency message dissemination in VANETs [6], [7], relay discovery in cooperative communication systems [1], [2], and sensor node selection in wireless sensor networks [5].

The paper is organized as follows. Section II describes the system model. The timer mappings when the prior distribution is known and unknown are developed in Sec. III and Sec. IV, respectively. Our conclusions follow in Sec. V.

II. System Model and Timer Scheme

Consider a system with k nodes and a sink. Neither the nodes nor the sink know k. Each node i maintains a metric $\mu_i \in \mathbb{R}^+$, which is not known to any other node, where $\mathbb{R}^+$ denotes the set of positive real numbers. The goal is for the sink to find the best node i^*, where

$$i^* = \underset{i \in \{1,2,\ldots,k\}}{\operatorname{argmax}} \ \mu_i. \tag{1}$$

We assume that the metrics are i.i.d. The independence assumption is justifiable as the metric depends on a local property of the node, such as the node's channel gain, which decorrelates with distance. The identicalness assumption ensures analytical tractability and is commonly made in the selection literature [1], [5], [13], [19], [20]. We shall investigate the case with non-identical metrics later in Sec. III-C3. Without loss of generality, we assume that the metrics are uniformly distributed in the interval $[0,1]$.[1]

A node i sets its timer T_i as a function of its metric μ_i as $T_i = f(\mu_i)$, where $f : [0,1] \to [0,\infty)$ is an MNI function. When the timer of a node expires, it transmits a *timer packet*. Furthermore, nodes whose timers expire after $T_{\max}$ do not transmit. The timer packet contains the identity of the node to enable the sink to identify which node transmitted. If two or more nodes transmit within a time window of Δ, a collision occurs and the sink cannot decode any of the transmissions. However, if only one node transmits, the receiver can decode the timer packet successfully [18, Chap. 4], [13], [19], [22]. This assumption is justified because the timer packet is a low payload packet, and its packet error rate can be made small by a conservative choice of the fading margin.

[1]This is because every node i can generate a new metric y_i as $y_i = F(\mu_i)$, where F is the cumulative distribution function (CDF) of the metric, which will be uniformly distributed over $[0,1]$. Since F is a monotone non-decreasing function, the node with the highest y_i is the same as the node with the highest μ_i. Assuming that the nodes know F is justifiable because a statistic such as the CDF varies at a time scale that is several orders of magnitude slower than that of the metric. The CDF can be estimated accurately using CDF estimation techniques [21].

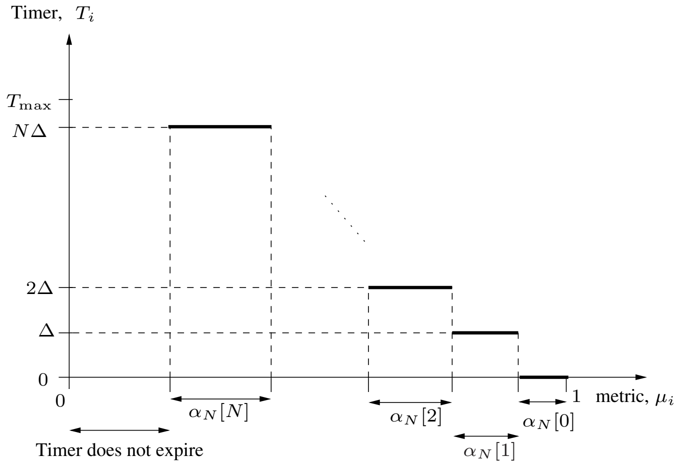

Fig. 1. Illustration of a staircase metric-to-timer mapping for a maximum selection duration of $T_{\max}$ and a vulnerability window Δ.

Motivated by the examples discussed earlier, we consider the following two models:

1) *Unknown number of nodes with known prior distribution:* In this model, the probability distribution of the number of nodes is known a priori. As discussed earlier, we consider the following two priors:

a) *Binomial prior*, in which the maximum possible number of nodes is K and the probability that the number of nodes is r is

$$\Pr[k = r] = \binom{K}{r} p^r (1-p)^{K-r}, \quad \text{for } 0 \le r \le K. \tag{2}$$

We shall refer to p as the participation probability.

b) *Poisson prior*, in which the average number of nodes is λ. Thus, the probability that the number of nodes is r is

$$\Pr[k = r] = e^{-\lambda}\frac{\lambda^r}{r!}, \quad \text{for } \ r \ge 0. \tag{3}$$

2) *Unknown number of nodes with unknown distribution:* In this case, all we are given is that k can be any number between $k_{\min}$ and $k_{\max}$, where $0 < k_{\min} \le k_{\max} < \infty$.[2]

III. Unknown Number of Nodes with Known Prior

Our goal is to find the optimal timer mapping that maximizes the average success probability. We first show that a timer mapping, of the form shown in Figure 1, is optimal. This result holds for any prior.

Theorem *1:* There exists an MNI timer mapping that maximizes the average success probability in which timers expire either only at $0, \Delta, \ldots, N\Delta$, or not at all, where $N = \left\lfloor \frac{T_{\max}}{\Delta} \right\rfloor$. When the metric μ lies in the interval $[1 - \alpha_N[0], 1)$, the timer expires immediately at time 0. When μ lies in $[1 - \alpha_N[0] - \alpha_N[1], 1 - \alpha_N[0])$, the timer expires at time Δ. In general, for $i = 0, 1, \ldots, N$, when $\mu \in \left[1 - \sum_{j=0}^{i} \alpha_N[j], 1 - \sum_{j=0}^{i-1} \alpha_N[j]\right)$, the timer expires at $i\Delta$. Timers of nodes whose metrics lie in the interval $\left[0, 1 - \sum_{j=0}^{N} \alpha_N[j]\right)$ do not expire at all.

Proof: The proof is presented in Appendix A. ■

The mapping looks like a staircase with the height of each stair being Δ and the length of the jth stair given by $\alpha_N[j]$. We, therefore, call it a *staircase mapping*. We shall refer to $\alpha_N[j]$ as the j^{th} *stair length* and N as the *number of timer levels*, which is completely determined by $T_{\max}$ and Δ. Thus, the problem reduces to optimizing the $N + 1$ stair lengths $\alpha_N[0], \alpha_N[1], \ldots, \alpha_N[N]$. We denote the $(N + 1)$-tuple $(\alpha_N[0], \alpha_N[1], \ldots, \alpha_N[N])$ by $\boldsymbol{\alpha}_N$. The staircase mapping is easy to implement in practice since each node needs to maintain a lookup table with $N + 1$ entries.

The above result generalizes the result in [11], which only proved that the staircase mapping is optimal when the number of nodes is known. Intuitively, this result is similar to the well known result in the MAC literature that slotted Aloha has a higher throughput than unslotted Aloha because the former reduces the probability of collisions by forcing the nodes to transmit only at the beginning of a time slot [18].

A. Optimal Timer Mapping for Binomial Prior

As shown in Appendix B, the average success probability $P_N(\boldsymbol{\alpha}_N)$ is given by

$$P_N(\boldsymbol{\alpha}_N) = Kp \sum_{i=0}^{N} \alpha_N[i] \left(1 - p\sum_{j=0}^{i} \alpha_N[j]\right)^{K-1}. \tag{4}$$

Therefore, the optimization problem can be stated as follows:

$$\mathcal{OB}: \qquad \underset{\boldsymbol{\alpha}_N}{\text{maximize}} \quad P_N(\boldsymbol{\alpha}_N) \tag{5}$$

$$\text{subject to} \quad \sum_{j=0}^{N} \alpha_N[j] \le 1, \tag{6}$$

$$\alpha_N[i] \ge 0; \ \text{for } 0 \le i \le N. \tag{7}$$

We now show that the $N + 1$ positivity constraints of (7) in problem $\mathcal{OB}$ are all inactive. Thus, the optimization problem is only constrained by (6). Let $\alpha_N^*[0], \alpha_N^*[1], \ldots, \alpha_N^*[N]$ denote the optimal stair lengths.

Lemma *1:* $\alpha_N^*[j] > 0$, for $j \in \{0, 1, \ldots, N\}$.

Proof: The proof is relegated to Appendix C. ■

The intuition behind this result is as follows. If any one of the $\alpha_N^*[j]$ were to be zero, then the N-level timer effectively reduces to a $(N - 1)$-level timer, and, hence, cannot do better than the optimal $(N - 1)$-level timer.

With the above result, we now present the complete solution to the problem $\mathcal{OB}$. It has a recursive structure.

Theorem *2:* Let $\boldsymbol{\beta}_N = (\beta_N[0], \beta_N[1], \ldots, \beta_N[N])$ be generated as follows:

$$\beta_N[i] = \begin{cases} \frac{1}{p}\left(\frac{1-P_{N-1}(\beta_{N-1}[0],\ldots,\beta_{N-1}[N-1])}{K-P_{N-1}(\beta_{N-1}[0],\ldots,\beta_{N-1}[N-1])}\right), & \text{if } i = 0 \\ (1 - p\beta_N[0])\beta_{N-1}[i-1], & \text{if } 1 \le i \le N \end{cases}, \tag{8}$$

where $\beta_0[0] = \frac{1}{Kp}$ and $P_N(\cdot)$ is given by (4).

If $\sum_{j=0}^{N} \beta_N[j] \le 1$ then $\boldsymbol{\alpha}_N^* = \boldsymbol{\beta}_N$. Otherwise,

$$\alpha_N^*[i] = \begin{cases} \frac{1}{p}\left(\frac{1-L_{N-1}^{\eta}(\boldsymbol{\alpha}_{N-1}^*)}{K-L_{N-1}^{\eta}(\boldsymbol{\alpha}_{N-1}^*)}\right), & \text{if } i = 0, \\ (1 - p\alpha_N^*[0])\alpha_{N-1}^*[i-1], & \text{if } 1 \le i \le N, \end{cases} \tag{9}$$

[2]Note that $k_{\min} = 0$ is not considered because, in this case, the worst case probability is always 0 and cannot be improved upon.

where

$$L_N^\eta(\boldsymbol{\alpha}_N^*) = P_N(\boldsymbol{\alpha}_N^*) + \eta \left(1 - p\sum_{j=0}^{N} \alpha_N^*[j]\right)^K, \tag{10}$$

and $\alpha_0^*[0] = \frac{1}{p}\left(\frac{1-\eta}{K-\eta}\right)$. Here, $\eta > 0$ is chosen such that $\sum_{j=0}^N \alpha_N^*[j] = 1$, and such a choice of η always exists.

Proof: The proof is given in Appendix D. ■

Comments: The recursion in (8) is similar to that in [11, (2)], except for the presence of the extra scaling factor p in our case. However, due to the constraint in (6), the stair lengths in [11] and $\alpha_N^*[j]$ need not be scaled versions of each other, except when p is small. Furthermore, η is found numerically, as is typical of several constrained optimization problems in wireless systems [3].

We see that N, K, and p together determine whether the constraint in (6) is active or not. The following lemma provides a sufficient condition for this. The utility of this result is that for this case, the optimal solution is given by $\boldsymbol{\beta}_N$ itself; the recursion in (9) is not required.

Lemma *2:* $\boldsymbol{\alpha}_N^* = \boldsymbol{\beta}_N$ if the average number of nodes Kp is greater than or equal to $N+1$.

Proof: The proof is relegated to Appendix E. ■

B. Optimal Timer Mapping with Poisson Prior

We now consider the case when the number of nodes k is a Poisson RV with mean λ. From Theorem 1, we know that there is an MNI staircase metric-to-timer mapping that is optimal. As shown in Appendix F, the average success probability is

$$P_N(\boldsymbol{\alpha}_N) = \lambda \sum_{i=0}^{N} \alpha_N[i] e^{-\lambda \sum_{j=0}^{i} \alpha_N[j]}. \tag{11}$$

Therefore, the average success probability maximization problem can be stated as

$$\mathcal{OP}: \quad \underset{\boldsymbol{\alpha}_N}{\text{maximize}} \quad P_N(\boldsymbol{\alpha}_N) \tag{12}$$

$$\text{subject to} \quad \sum_{j=0}^{N} \alpha_N[j] \le 1, \tag{13}$$

$$\alpha_N[i] \ge 0; \ \text{for } 0 \le i \le N. \tag{14}$$

Let $\boldsymbol{\alpha}_N^* = (\alpha_N^*[0], \ldots, \alpha_N^*[N])$ denote the solution of problem $\mathcal{OP}$. Then, similar to Lemma 1 for the binomial case, the following lemma shows that none of the stair lengths can be zero.

Lemma *3:* $\alpha_N^*[j] > 0$, for all $j \in \{0, 1, \ldots, N\}$.

Proof: The proof is similar to that in Appendix C and is not repeated here to conserve space. ■

The intuition behind this result is the same as that for Lemma 1. With the above result, the following theorem provides the solution to the problem $\mathcal{OP}$.

Theorem *3:* Let $\boldsymbol{\beta}_N = (\beta_N[0], \beta_N[1], \ldots, \beta_N[N])$ be generated as follows:

$$\beta_N[i] = \begin{cases} \frac{1 - P_{N-1}(\beta_{N-1}[0], \ldots, \beta_{N-1}[N-1])}{\lambda}, & \text{if } i = 0, \\ \beta_{N-1}[i-1], & \text{if } 1 \le i \le N, \end{cases} \tag{15}$$

where $\beta_0[0] = \frac{1}{\lambda}$ and $P_N(\cdot)$ is given by (11).

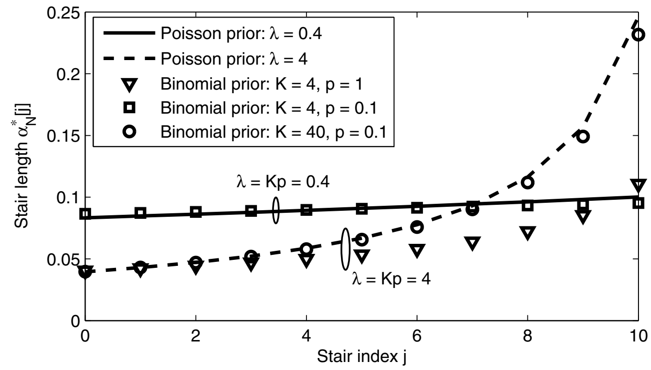

Fig. 2. Poisson and binomial priors: Optimal stair lengths for different values of K, p, and λ ($N = 10$).

If $\sum_{j=0}^N \beta_N[j] \le 1$ then $\boldsymbol{\alpha}_N^* = \boldsymbol{\beta}_N$. Otherwise,

$$\alpha_N^*[i] = \begin{cases} \frac{1 - H_{N-1}^\gamma(\boldsymbol{\alpha}_{N-1}^*)}{\lambda}, & \text{if } i = 0, \\ \alpha_{N-1}^*[i-1], & \text{if } 1 \le i \le N, \end{cases} \tag{16}$$

where

$$H_N^\gamma(\boldsymbol{\alpha}_N^*) = P_N(\boldsymbol{\alpha}_N^*) + \gamma e^{-\lambda \sum_{j=0}^N \alpha_N^*[j]}, \tag{17}$$

and $\alpha_0^*[0] = \frac{1-\gamma}{\lambda}$. Here, $\gamma > 0$ is chosen such that $\sum_{j=0}^N \alpha_N^*[j] = 1$, and such a choice of γ always exists.

Proof: The proof is relegated to Appendix G. ■

Note that for $\gamma = 0$, the recursion in (16) reduces to that in (15). As in the binomial case, γ is found numerically.

A sufficient condition for $\boldsymbol{\beta}_N$, which is analogous to Lemma 2, to be the solution of $\mathcal{OP}$ is as follows.

Lemma *4:* $\boldsymbol{\alpha}_N^* = \boldsymbol{\beta}_N$ if $\lambda \ge N+1$.

Proof: The proof is relegated to Appendix H. ■

Determining the optimal stair lengths for a general prior distribution is an open problem. However, the Poisson prior design can be applied to a more general class of distributions using the Brun's sieve result [23, Chap. 10].

C. Numerical Results

We now present numerical results to better understand the optimal mapping and also to benchmark its performance.

1) Optimal Stair Lengths: Figure 2 plots the optimal stair lengths as a function of the stair index j for various λ for the Poisson prior and K and p for the binomial prior. We observe here that for larger λ or Kp the stair lengths increase as j increases. However, for smaller values of λ or Kp the stair lengths become almost equal. When $\lambda = Kp$ we see that for smaller stair indices, the stair lengths of the two priors are almost equal. However, for larger indices, the stair lengths are larger for the Poisson prior than for the binomial prior. Further, notice that as $K \to \infty$ and $p \to 0$, with $Kp = \lambda$, the two stair lengths tend to be equal. This is intuitive because, in this case, the binomial distribution converges to the Poisson distribution [24, Chap. 2].

2) Benchmarks and Performance Comparisons: To better understand the performance of the optimal mapping, we benchmark it against four designs: (i) *Design for K nodes:* Here, the stair lengths are designed assuming K nodes are present in the system always, i.e., by setting $p = 1$ in

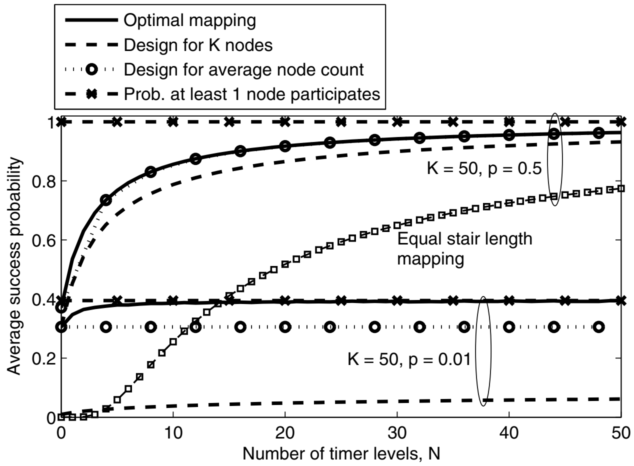

Fig. 3. Binomial prior: Comparison of average success probability vs. the number of timer levels ($K = 50$).

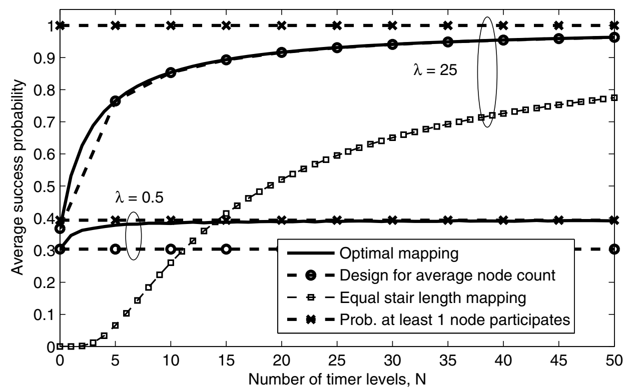

Fig. 4. Poisson prior: Benchmarking of average success probability as a function of the number of timer levels for optimal mapping, design for average node count, and equal stair length mapping for $\lambda = 0.5$ and $\lambda = 25$.

Theorem 2. (ii) *Design for average node count:* Here, the stair lengths are instead designed assuming that the number of nodes in the system is equal to its average value. (iii) *Equal stair length mapping [13]:* Here, all the stair lengths are set to $\frac{1}{N+1}$. (iv) *Inverse mapping [1]:* Here, the mapping is given by $f(\mu) = c/\mu$, where $c > 0$ is a constant and is optimized numerically to maximize the average success probability. We also plot the probability that at least one node is present in the system; this equals $1 - (1-p)^K$ for the binomial prior and $1 - e^{-\lambda}$ for the Poisson prior. It is an upper bound on the selection probability because any selection is bound to fail if there are no nodes in the system.

Figure 3 plots the average success probability of the various timer mappings for the binomial prior as a function of the number of timer levels N. We see that the average success probability of the optimal mapping increases with N, and is very close to the upper bound. For example, for $N = 20$ and $p = 0.01$, it is 99% of the upper bound. Further, we see that for $p = 0.5$, the design for average node count performs almost as well as the optimal mapping. However, when p is small, e.g., $p = 0.01$, the design for average node count has a 29% lower average success probability. The optimal mapping achieves 6% more average success probability than the design for K nodes for $p = 0.5$ when $N = 20$; for $p = 0.01$ this increases markedly to 715%. While the equal stair length mapping is close to optimal for small p, as we saw already in Figure 2, for $p = 0.5$, its average success probability is 44% lower than that of the optimal mapping. The average success probability of the inverse mapping is not plotted in order to avoid clutter. Its performance is the worst among all schemes. For example, when $N = 20$ and $p = 0.5$, its average success probability is 0.37, while that of the optimal mapping is 0.92. We, therefore, do not consider it in the rest of this section. Figure 4 plots the corresponding results for the Poisson prior. The trends are similar to those in Figure 3.

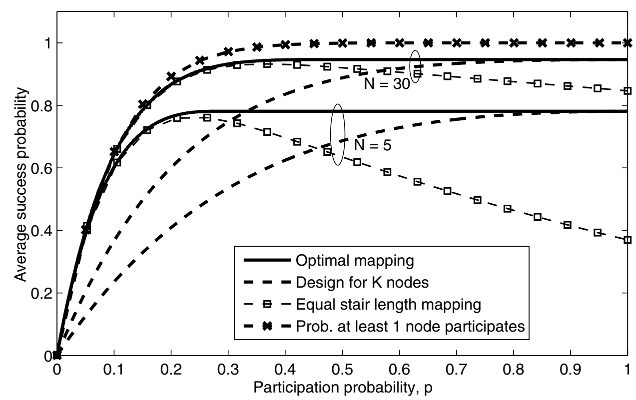

Fig. 5. Binomial prior: Comparison of average success probability vs. the participation probability p for different values of N ($K = 10$).

Figure 5 plots the average success probability for the binomial prior as a function of the participation probability p for $N = 5$ and $N = 10$. For the optimal mapping, we observe that the average success probability is very close to its upper bound when p is small and is, thus, limited by the absence of nodes to select form. Consequently, increasing N for $p < 0.1$ does not increase the average success probability. However, for larger p, this trend changes. For example, when $p = 0.6$, we see that the optimal mapping's average success probability increases by 21% when N is increased from 5 to 30. The trends for the average success probability as a function of λ for the Poisson prior are qualitatively similar.

3) Non-identical Metrics: To understand the impact of the non-identicalness of the metrics, we consider below a wireless system with K nodes and a sink, in which a node has data to transmit with probability p. Thus, it participates in the selection process with probability p. To model the fact that metrics of the nodes are not identical, we set the mean of the channel power gain h_i of node i to be $h_0\theta^{i-1}$, for $i = 1, 2, \ldots, K$. The further θ is away from unity, the more statistically non-identical the metrics are. The channels are mutually independent, and undergo Rayleigh fading. The objective is to select the node with the highest channel power gain. A node i sets is metric μ_i as $\mu_i = 1 - e^{-\frac{h_i}{h_0}}$. This ensures that the metrics of all the nodes lie between 0 and 1, as required by the proposed scheme.[3] In this case, the average

[3]It requires the nodes to know the largest mean channel power gain. Since this is a statistic, it changes at a rate that is orders of magnitude slower than the instantaneous channel gains. It can, therefore, be estimated by the sink and communicated to all the nodes.

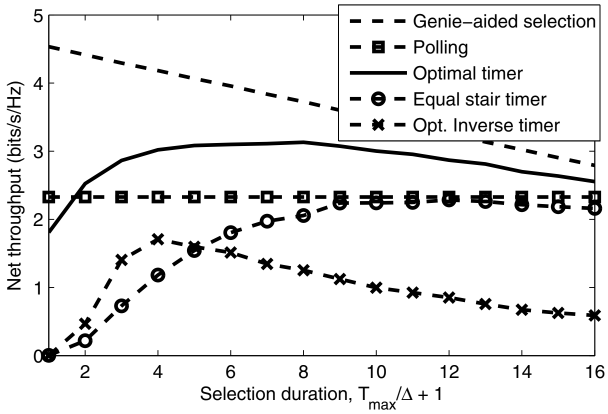

Fig. 6. Net throughput as a function of selection duration ($K = 20$, $p = 0.4$, and $T_c = 40\Delta$).

success probability, in fact, increases marginally from 0.749 for $\theta = 1$ to 0.753 for $\theta = 0.9$ and to 0.8 for $\theta = 0.1$. This is because the above transformation ensures that the distribution of a non-best node's metric is likely to be skewed towards 0 than 1, which reduces the collision probability.

4) Net Throughput Comparisons: We now study the impact of the selection scheme on the net throughput of the system. This performance measure captures the time spent on selection, the impact of reliability of the selection process in selecting the best node, and the benefits obtained from opportunistic selection.

As in Sec. III-C3, we consider a system with K nodes and a sink, in which each node has data to transmit with probability p. The channel power gains remain constant for a coherence time T_c and change independently thereafter. The channel power gains of the nodes are i.i.d. and are exponentially distributed with a mean of 10 dB. The sink needs to select the node with the highest channel power gain and then transmits data to it. The selection phase runs for a duration of $T_{\max}+\Delta$ and is immediately followed by the data transmission phase, which lasts for a duration of $T_c - T_{\max} - \Delta$.[4] For illustration purposes, the data transmission rate is $\log_2(1 + h_{\text{sel}})$, where h_{sel} is the channel power gain of the selected node.

Figure 6 compares the net throughput of several timer-based selection schemes. Also plotted are the net throughputs of genie-aided selection and polling:

1) *Genie-aided selection:* In this, after the selection duration of $T_{\max} + \Delta$, the sink is told by a genie who the best node is. This serves as an upper bound for all the timer-based selection schemes.
2) *Polling:* In this, all the K nodes sequentially reveal their metrics to the sink. This takes a duration of $K\Delta$ to select.

We see that the proposed scheme markedly outperforms the other schemes except when $N \leq 1$. For a selection duration of 8Δ, the proposed scheme achieves 35%, 52%, and 150% higher net throughput than polling, equal stair length timer, and optimized inverse timer, respectively. Notice that there exists an optimal choice of selection duration at which the net throughput is maximized.

[4]The selection duration is $T_{\max} + \Delta$ because a timer, in the worst case, can expire at $T_{\max}$. A Δ duration after this is allowed so that the sink can receive this transmission.

IV. Unknown Number of Nodes with Unknown Distribution

We now consider the general case where even the probability distribution of the number of nodes is unknown. All we know is that k lies between $k_{\min}$ and $k_{\max}$, where $0 < k_{\min} \leq k_{\max} < \infty$. Our goal is to arrive at a robust design that maximizes the worst case average success probability over all possible probability distributions. The approach in Sec. III can be interpreted as a softer version of the formulation pursued here because the success probability with k nodes is weighted by the probability that there are k nodes in the system.

Let $\mathcal{P}$ be the set of all probability distributions on the set $\mathcal{K} = \{k_{\min}, k_{\min}+1, \ldots, k_{\max}\}$. For each distribution $Q \in \mathcal{P}$, let $Q(i)$ denote the probability that the number of nodes is i. Further, $Q(i) = 0$ if $i \notin \mathcal{K}$. Let $S_i(f)$ be the success probability when i nodes contend with timer mapping f. Then the average success probability, given a distribution Q on the number of nodes, which we denote by $\mathbb{E}_Q[S_i(f)]$, is given by

$$\mathbb{E}_Q[S_i(f)] = \sum_{i\in\mathcal{K}} Q(i) S_i(f). \tag{18}$$

The optimization problem can be stated as follows:

$$\mathcal{OU}: \quad \underset{f \text{ is MNI}}{\text{maximize}} \quad \underset{Q\in\mathcal{P}}{\text{minimize}} \quad \mathbb{E}_Q[S_i(f)]. \tag{19}$$

The following result shows that it is sufficient to search for the optimal mapping over the space of only $(k_{\max} - k_{\min} + 1)$ point mass distributions instead of over the set $\mathcal{P}$, whose size is uncountably infinite.

Lemma *5:* The two problems $\text{minimize}_{Q\in\mathcal{P}}\ \mathbb{E}_Q[S_i(f)]$ and $\text{minimize}_{i\in\mathcal{K}}\ S_i(f)$ are equivalent, i.e.,

$$\underset{Q\in\mathcal{P}}{\text{minimize}}\ \mathbb{E}_Q[S_i(f)] = \underset{i\in\mathcal{K}}{\text{minimize}}\ S_i(f). \tag{20}$$

Proof: The proof is relegated to Appendix I. ■

The following theorem shows that the optimal mapping is again a staircase mapping.

Theorem *4:* The MNI staircase mapping in which the timers expire either at $0, \Delta, 2\Delta, \ldots, N\Delta$ or not at all, solves $\mathcal{OU}$.

Proof: The proof is relegated to Appendix J. ■

As in Sec. III, the stair lengths $\boldsymbol{\alpha}_N = (\alpha_N[0], \ldots, \alpha_N[N])$ completely determine such a staircase mapping. Thus, it suffices to optimize $\boldsymbol{\alpha}_N$. We denote the success probability when k nodes contend as $\Lambda_{N,k}(\boldsymbol{\alpha}_N)$, which from Appendix B is given by

$$\Lambda_{N,k}(\boldsymbol{\alpha}_N) = \sum_{i=0}^{N} k\alpha_N[i] \left(1 - \sum_{j=0}^{i} \alpha_N[j]\right)^{k-1}. \tag{21}$$

Therefore, the optimization problem can be restated as:

$$\mathcal{OU}': \quad \underset{\boldsymbol{\alpha}_N}{\text{maximize}} \quad \underset{k\in\mathcal{K}}{\text{minimize}} \quad \Lambda_{N,k}(\boldsymbol{\alpha}_N), \tag{22}$$

$$\text{subject to} \quad \sum_{j=0}^{N} \alpha_N[j] \leq 1, \tag{23}$$

$$\alpha_N[j] \geq 0. \tag{24}$$

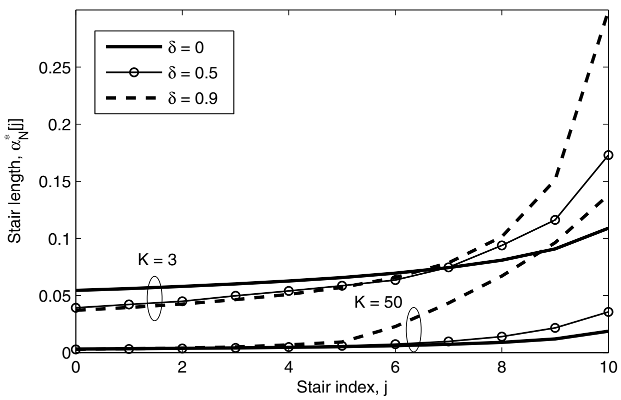

Fig. 7. Stair lengths of the robust mapping as a function of j for different uncertainties.

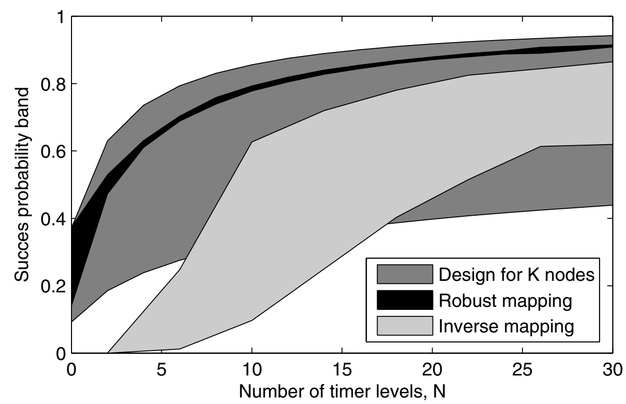

Fig. 8. Comparison of success probability bands as a function of the number of timer levels of the robust mapping, inverse mapping, and design for K nodes ($K = 50$ and $\delta = 0.9$).

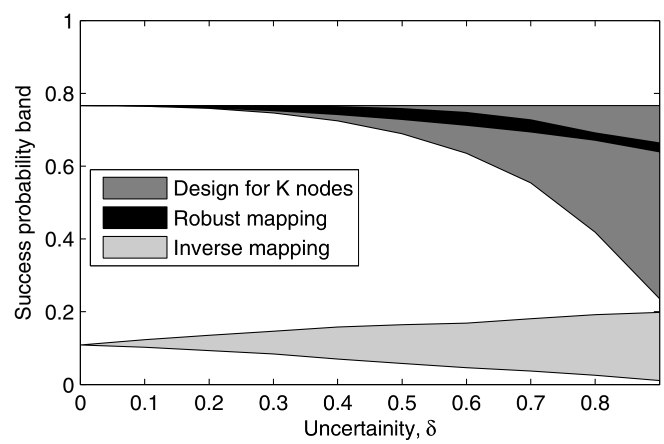

Fig. 9. Effect of uncertainty on success probability band of various timer schemes ($K = 100$ and $N = 5$).

Since $\Lambda_{N,k_{\min}}(\boldsymbol{\alpha}_N), \Lambda_{N,k_{\min}+1}(\boldsymbol{\alpha}_N), \ldots, \Lambda_{N,k_{\max}}(\boldsymbol{\alpha}_N)$ are continuous functions in $\boldsymbol{\alpha}_N$, the function $\min_{k\in\mathcal{K}} \Lambda_{N,k}(\boldsymbol{\alpha}_N)$ is also a continuous function in $\boldsymbol{\alpha}_N$. Further, the constraint region of $\boldsymbol{\alpha}_N$, defined by (23) and (24), is closed and bounded. Since a continuous function over a closed bounded set always attains its maximum [25], there exists an $\boldsymbol{\alpha}_N^* = (\alpha_N^*[0], \alpha_N^*[1], \ldots, \alpha_N^*[N])$ that lies in the constraint region that solves the problem $\mathcal{OU}'$. However, the function $G(\boldsymbol{\alpha}_N) = \min_{k\in\mathcal{K}} \Lambda_{N,k}(\boldsymbol{\alpha}_N)$ is a non-convex and non-differentiable function.

We, therefore, use the `fmincon` function of MATLAB, which uses sequential quadratic programming, to numerically solve $\mathcal{OU}'$. We have observed that it always solves the optimization problem. The reason why it does so is that $G(\boldsymbol{\alpha}_N)$, although non-differentiable on the entire constraint set, is infinitely differentiable over a dense subset of the constraint set.[5] Secondly, if the algorithm reaches a non-differentiable point, then it computes a second derivative, assuming the function to be smooth, and uses it to proceed further.

A. Numerical Results

To build intuition, let $k_{\min} = (1-\delta)K$ and $k_{\max} = (1+\delta)K$, where δ characterizes the uncertainty in the knowledge of the number of nodes. We shall refer to δ as the *uncertainty*.

Figure 7 plots the stair lengths of the robust mapping for different values of δ and K for $N = 9$. We observe that $\alpha_N^*[j]$ increases as j increases. Thus, the mapping gets more aggressive in making nodes transmit when the time that remains for selection decreases, as was also the case in Sec. III for the known prior case. Further, as δ increases or K decreases, the stair length $\alpha_N^*[j]$ of the robust mapping increases as j increases.

In order to better visualize and compare the performance of the timer schemes, we plot the success probability band, which is the range of values of the average success probability, of each scheme. Figure 8 plots the success probability bands for: (i) design for K nodes, (ii) inverse mapping, and (iii) robust mapping. Notice that the success probability band for the robust mapping is significantly narrower than all the other schemes, which shows its robustness to uncertainty in the number of nodes. The inverse mapping fares poorly compared to the robust mapping, despite its parameters being separately optimized for each δ. Another key observation is that the worst case success probability of the robust mapping increases as N increases, and is markedly better than all the other schemes.

Figure 9 plots the success probability bands as a function of δ. We observe that the success probability band of the inverse mapping and the design for K nodes widens appreciably as δ increases. However, for the robust mapping, it remains narrow. The equal stair mapping is not shown to avoid clutter. It is the least robust to uncertainty among all the above schemes. For example, when $\delta = 0.9$ and $N = 20$, its success probability varies from 0.05 to 0.9. Further, for smaller N, its worst case success probability is almost zero, even when δ is as small as 0.3.

V. Conclusions

The probability with which the popular, distributed timer-based selection scheme selects the best node depends on the monotonically non-increasing metric-to-timer mapping it uses. We developed optimal timer mappings for the practical scenario in which the number of nodes in the system is not known. Two models for unknown number of nodes, namely,

[5]A subset A of a set S is called *dense* if every point of S either belongs to A or is a limit point of A [25].

known prior and unknown prior distribution, were considered. For both models, we saw that an optimal timer mapping is a staircase mapping in which the timers expire either at $0, \Delta, \ldots, N\Delta$ or not at all.

For the binomial and Poisson priors, which arise in several systems, we showed that the optimal stair lengths can be computed using a recursion in the number of timer levels. This is unlike the ad hoc mappings proposed in the literature, which strive to provide an explicit functional form for the mapping. For the unknown prior case, the proposed robust mapping achieves the highest worst case success probability among all the mappings and has the narrowest success probability band.

APPENDIX

A. Proof of Theorem 1

Let $f : [0, 1] \to [0, \infty)$ be an MNI mapping. Define a new mapping g as follows:

$$g(\mu) = \begin{cases} \left\lfloor \frac{f(\mu)}{\Delta} \right\rfloor \Delta, & \text{if } f(\mu) \leq T_{\max}, \\ T_{\max} + \epsilon, & \text{if } f(\mu) > T_{\max}, \end{cases} \quad (25)$$

where $\epsilon > 0$. This ensures that a node does not transmit if its timer expires after $T_{\max}$. This new mapping g can be shown to be an MNI mapping because f is MNI. Further, g has the same structure as described in Theorem 1. We show below that for any $k \in \mathbb{N}$ and for any realization of the metrics $\zeta_1, \ldots, \zeta_k$, using g results in a success if using f results in a success. Thus, the average success probability of g is greater than or equal to that of f.

If $k = 0$, then the success probability is 0 for both g and f since there are no nodes in the system. If $k = 1$ and the user's metric is ζ_1, then from (25), $g(\zeta_1) \leq T_{\max}$ if $f(\zeta_1) \leq T_{\max}$. Thus, g will result in a success whenever f results in a success.

Now, consider the last case where $k \geq 2$. Let $[i]$ denote the node with the ith highest metric among the k nodes. There are only two cases in which f succeeds: (i) *When* $f(\zeta_{[2]}) > T_{\max}$ *and* $f(\zeta_{[1]}) \leq T_{\max}$*:* In this case, from (25), $g(\zeta_{[1]}) \leq T_{\max}$ and $g(\zeta_{[j]}) = T_{\max} + \epsilon$, for all $j \neq 1$. Thus, only the best node transmits its timer packet even with g and a success will occur. (ii) *When* $f(\zeta_{[2]}) \leq T_{\max}$ *and* $|f(\zeta_{[1]}) - f(\zeta_{[2]})| > \Delta$*:* In this case, from (25), we get $g(\zeta_{[2]}) \leq T_{\max}$. Furthermore, from the property of the floor function, $|g(\zeta_{[1]}) - g(\zeta_{[2]})| > \Delta$. Thus, even in this case, a success will occur and g will select the best node.

B. Derivation of (4)

Let $\Lambda_{N,k}(\boldsymbol{\alpha}_N)$ denote the success probability when k nodes participate in the selection process. Recall from Appendix A that $\mu_{[1]} \geq \mu_{[2]} \geq \cdots \geq \mu_{[k]}$. Summing over the mutually exclusive events in which the best node's metric lies in the ith interval, for $0 \leq i \leq N$, we get

$$\Lambda_{N,k}(\boldsymbol{\alpha}_N) = \sum_{i=0}^{N} \Pr\left[\mu_{[1]} \in \left[1 - \sum_{j=0}^{i} \alpha_N[j], 1 - \sum_{j=0}^{i-1} \alpha_N[j]\right) \right. \\ \left. \text{and } \mu_{[2]} \leq 1 - \sum_{j=0}^{i} \alpha_N[j]\right].$$

Since the metrics $\mu_1, \ldots, \mu_k$ are i.i.d. and $\mu_{[k]} \leq \cdots \leq \mu_{[2]}$, we get

$$\begin{aligned} &\Lambda_{N,k}(\boldsymbol{\alpha}_N) \\ &= \sum_{i=0}^{N} k \Pr\left[\mu_1 \in \left[1 - \sum_{j=0}^{i} \alpha_N[j], 1 - \sum_{j=0}^{i-1} \alpha_N[j]\right) \right. \\ &\qquad \left. \text{and } \mu_l \leq 1 - \sum_{j=0}^{i} \alpha_N[j], \text{ for } 2 \leq l \leq k\right], \\ &= \sum_{i=0}^{N} k \alpha_N[i] \left(1 - \sum_{j=0}^{i} \alpha_N[j]\right)^{k-1}. \end{aligned} \quad (26)$$

Since k is a binomial RV, the average success probability $P_N(\boldsymbol{\alpha}_N)$ is given by

$$P_N(\boldsymbol{\alpha}_N) = \sum_{k=0}^{K} \binom{K}{k} p^k (1-p)^{K-k} \Lambda_{N,k}(\boldsymbol{\alpha}_N). \quad (27)$$

Substituting (26) yields

$$P_N(\boldsymbol{\alpha}_N) = \sum_{k=0}^{K} \binom{K}{k} p^k (1-p)^{K-k} \sum_{i=0}^{N} k \alpha_N[i] \\ \times \left(1 - \sum_{j=0}^{i} \alpha_N[j]\right)^{k-1}. \quad (28)$$

Interchanging the two summations and using the binomial expansion yields (4).

C. Proof of Lemma 1

Let there be a $j \in \{0, 1, \ldots, N\}$ such that $\alpha_N^*[j] = 0$. Since the metrics are uniformly distributed over $[0, 1]$, the success probability of the mapping with stair lengths $\alpha_N^*[0], \ldots, \alpha_N^*[N]$ would be same as that of a mapping with stair lengths $\alpha_N[0], \ldots, \alpha_N[N]$, where $\alpha_N[i] = \alpha_N^*[i]$, for $i < j$, $\alpha_N[i] = \alpha_N^*[i+1]$, for $j \leq i < N$, and $\alpha_N[N] = 0$. Thus, it suffices to argue that $\alpha_N^*[N]$ cannot be equal to 0. This is easy to see because, for an ϵ_1 that is sufficiently small,

$$P_N(\alpha_N^*[0], \ldots, \alpha_N^*[N-1], \epsilon_1) > P_N(\alpha_N^*[0], \ldots, \alpha_N^*[N-1], 0).$$

This follows because $P_N(\alpha_N[0], \ldots, \alpha_N[N])$ is an increasing function of $\alpha_N[N]$ at $\alpha_N[N] = 0$. The right hand partial derivative of $P_N(\alpha_N[0], \ldots, \alpha_N[N])$ with respect to $\alpha_N[N]$ at $\alpha_N[N] = 0$ is given by

$$\left.\frac{\partial P_N(\alpha_N[0], \ldots, \alpha_N[N])}{\partial \alpha_N[N]_+}\right|_{\alpha_N[N]=0} \\ = \left.\frac{\partial K p \sum_{i=0}^{N} \alpha_N[i] \left(1 - p \sum_{j=0}^{i} \alpha_N[j]\right)^{K-1}}{\partial \alpha_N[N]}\right|_{\alpha_N[N]=0}.$$

This equals $Kp\left(1 - p\sum_{i=0}^{N-1} \alpha_N[i]\right)^{K-1}$, which is positive when $\sum_{i=0}^{N-1} \alpha_N[i] \leq 1$.

D. Proof of Theorem 2

Define an auxiliary function $L_N^\eta(\boldsymbol{\alpha}_N)$ as

$$L_N^\eta(\boldsymbol{\alpha}_N) = P_N(\boldsymbol{\alpha}_N) + \eta\left(1 - p\sum_{j=0}^{N}\alpha_N[j]\right)^K, \quad (29)$$

where $\eta \geq 0$. Here, $\left(1 - p\sum_{j=0}^{N}\alpha_N[j]\right)^K$ can be interpreted as a displeasure function that characterizes the displeasure or penalty when the constraint $\sum_{j=0}^{N}\alpha_N[j] \leq 1$ is not met. Physically, it is the average probability that no node's timer expires; we shall, therefore, call it the *average idle probability*. In the first part of the proof, we derive the optimal $\tilde{\boldsymbol{\alpha}}_N$ that maximizes $L_N^\eta(\boldsymbol{\alpha}_N)$. In the second part, we show that $\tilde{\boldsymbol{\alpha}}_N$ also solves $\mathcal{OB}$ for an appropriate choice of η.

Derivation of $\tilde{\boldsymbol{\alpha}}_N$: $L_N^\eta(\boldsymbol{\alpha}_N)$ can be shown to be equal to

$$L_N^\eta(\boldsymbol{\alpha}_N) = Kp\alpha_N[0]\left(1 - p\alpha_N[0]\right)^{K-1}$$
$$+Kp\sum_{i=1}^{N}\alpha_N[i]\left(1 - p\sum_{j=0}^{i}\alpha_N[j]\right)^{K-1} + \eta\left(1 - p\sum_{j=0}^{N}\alpha_N[j]\right)^K.$$

Taking $\left(1 - p\alpha_N[0]\right)^K$ as a common factor from the last two terms, we get

$$L_N^\eta(\boldsymbol{\alpha}_N) = Kp\alpha_N[0]\left(1 - p\alpha_N[0]\right)^{K-1} + \left(1 - p\alpha_N[0]\right)^K \times L_{N-1}^\eta\left(\frac{\alpha_N[1]}{1 - p\alpha_N[0]}, \ldots, \frac{\alpha_N[N]}{1 - p\alpha_N[0]}\right). \quad (30)$$

Thus,

$$L_N^\eta(\boldsymbol{\alpha}_N) \leq Kp\alpha_N[0]\left(1 - p\alpha_N[0]\right)^{K-1} + \left(1 - p\alpha_N[0]\right)^K L_{N-1}^\eta(\tilde{\boldsymbol{\alpha}}_{N-1}), \quad (31)$$

where $\tilde{\boldsymbol{\alpha}}_{N-1} = (\tilde{\alpha}_{N-1}[0], \ldots, \tilde{\alpha}_{N-1}[N-1])$ maximizes L_{N-1}^η. Furthermore, from (30) and (31), given any $\alpha_N[0] \in (0, 1)$, the upper bound in (31) can indeed be achieved by setting

$$\alpha_N[j] = \left(1 - p\alpha_N[0]\right)\tilde{\alpha}_{N-1}[j-1], \quad \text{for } 1 \leq j \leq N. \quad (32)$$

Hence, we have

$$\max_{\boldsymbol{\alpha}_N} L_N^\eta(\boldsymbol{\alpha}_N) = \max_{\alpha_N[0]\in(0,1)}\left\{Kp\alpha_N[0]\left(1 - p\alpha_N[0]\right)^{K-1} + \left(1 - p\alpha_N[0]\right)^K L_{N-1}^\eta(\tilde{\boldsymbol{\alpha}}_{N-1})\right\}. \quad (33)$$

Using the first order condition, this maximum is achieved when $\alpha_N[0] = \frac{1}{p}\left(\frac{1 - L_{N-1}^\eta(\tilde{\boldsymbol{\alpha}}_{N-1})}{K - L_{N-1}^\eta(\tilde{\boldsymbol{\alpha}}_{N-1})}\right)$. Thus, using (32), we see that $\tilde{\boldsymbol{\alpha}}_N$ is given by

$$\tilde{\alpha}_N[i] = \begin{cases} \frac{1}{p}\left(\frac{1 - L_{N-1}^\eta(\tilde{\boldsymbol{\alpha}}_{N-1})}{K - L_{N-1}^\eta(\tilde{\boldsymbol{\alpha}}_{N-1})}\right), & \text{if } i = 0, \\ \left(1 - p\tilde{\alpha}_N[0]\right)\tilde{\alpha}_{N-1}[i-1], & \text{if } 1 \leq i \leq N. \end{cases} \quad (34)$$

For $N = 0$, it can be easily shown that $\tilde{\alpha}_0[0] = \frac{1}{p}\left(\frac{1-\eta}{K-\eta}\right)$ maximizes

$$L_0^\eta(\alpha_0[0]) = Kp\alpha_0[0]\left(1 - p\alpha_0[0]\right)^{K-1} + \eta\left(1 - p\alpha_0[0]\right)^K.$$

We now investigate the solution for two special values of η, namely, 0 and 1.

a) *$\tilde{\boldsymbol{\alpha}}_N$ for $\eta = 0$:* For $\eta = 0$, $L_N^\eta(\boldsymbol{\alpha}_N) = P_N(\boldsymbol{\alpha}_N)$ and the recursion in (34) and its initial condition are the same as that in (8). Thus, $\tilde{\boldsymbol{\alpha}}_N = \boldsymbol{\beta}_N$, where $\boldsymbol{\beta}_N$ is given by (8).

b) *$\tilde{\boldsymbol{\alpha}}_N$ for $\eta = 1$:* For $\eta = 1$, $L_N^\eta(\boldsymbol{\alpha}_N)$ is given by

$$L_N^\eta(\boldsymbol{\alpha}_N) = P_N(\boldsymbol{\alpha}_N) + \left(1 - p\sum_{j=0}^{N}\alpha_N[j]\right)^K. \quad (35)$$

This is nothing but the sum of the success and idle probabilities because $\left(1 - p\sum_{j=0}^{N}\alpha_N[j]\right)^K$ is the probability that no node's timer expires; thus, it is upper bounded by 1. This upper bound is achieved when $\tilde{\alpha}_N[i] = 0$, for all $i \in \{0, 1, \ldots, N\}$. Hence, for $\eta = 1$, $\sum_{j=0}^{N}\tilde{\alpha}_N[j] = 0$.

Note: The unconstrained maximization of $L_N^\eta(\boldsymbol{\alpha}_N)$ can also be stated and solved as a $(N+1)$-horizon dynamic programming problem [26] whose action space at time i is the stair length $\alpha_N[i-1]$, for $1 \leq i \leq N+1$. Its state space is the entire metric interval, and the state at time i is $1 - p\sum_{j=0}^{N}\alpha_N[j]$, for $1 \leq i \leq N+1$, and (33) can be interpreted as the Bellman equation.

Proof of optimality of $\boldsymbol{\alpha}_N^*$: If $\sum_{j=0}^{N}\beta_N[j] \leq 1$, then $\boldsymbol{\beta}_N$ is feasible and solves problem $\mathcal{OB}$. Hence, $\boldsymbol{\alpha}_N^* = \boldsymbol{\beta}_N$.

Consider now the case where $\sum_{j=0}^{N}\beta_N[j] > 1$. Then, from the intermediate value theorem [25], there is an $\eta \in (0, 1)$ such that $\sum_{j=0}^{N}\tilde{\alpha}_N[j] = 1$.[6] Clearly, such an $\tilde{\boldsymbol{\alpha}}_N$ is also feasible. Further, for this η, the auxiliary function is given by

$$L_N^\eta(\tilde{\boldsymbol{\alpha}}_N) = P_N(\tilde{\boldsymbol{\alpha}}_N) + \eta\left(1 - p\right)^K. \quad (36)$$

By definition, for any feasible $\boldsymbol{\alpha}_N$, we have $L_N^\eta(\tilde{\boldsymbol{\alpha}}_N) \geq L_N^\eta(\boldsymbol{\alpha}_N)$. From (29), this implies that

$$P_N(\tilde{\boldsymbol{\alpha}}_N) \geq P_N(\boldsymbol{\alpha}_N) + \eta\left[\left(1 - p\sum_{j=0}^{N}\alpha_N[j]\right)^K - \left(1 - p\right)^K\right]. \quad (37)$$

Since $\sum_{j=0}^{N}\alpha_N[j] \leq 1$ and $\eta \geq 0$, it follows that $\eta\left[\left(1 - p\sum_{j=0}^{N}\alpha_N[j]\right)^K - \left(1 - p\right)^K\right] \geq 0$. Hence, for every feasible $\boldsymbol{\alpha}_N$, $P_N(\tilde{\boldsymbol{\alpha}}_N) \geq P_N(\boldsymbol{\alpha}_N)$. Therefore, $\boldsymbol{\alpha}_N^* = \tilde{\boldsymbol{\alpha}}_N$.

E. Proof of Lemma 2

We first show that $\beta_N[j] \leq \frac{1}{Kp}$, for $0 \leq j \leq N$. From (8),

$$\beta_N[0] = \frac{1}{p}\left(\frac{1 - P_{N-1}(\boldsymbol{\beta}_{N-1})}{K - P_{N-1}(\boldsymbol{\beta}_{N-1})}\right). \quad (38)$$

Since $P_{N-1}(\boldsymbol{\beta}_{N-1}) \geq 0$, we get $\beta_N[0] \leq \frac{1}{Kp}$. Applying the recursion in (8) repeatedly, we get

$$\beta_N[j] = \beta_{N-j}[0]\prod_{i=0}^{j-1}\left(1 - p\beta_{N-i}[0]\right) \leq \beta_{N-j}[0] \leq \frac{1}{Kp}. \quad (39)$$

[6] To apply the intermediate value theorem, we need to show that $\tilde{\alpha}_N[j]$, for $j = 0, 1, \ldots, N$, are continuous functions in η. We use induction on N to establish this. First, for $N = 0$, $\tilde{\alpha}_0[0] = \frac{1}{p}\left(\frac{1-\eta}{K-\eta}\right)$, which is a continuous function in η, for $0 \leq \eta \leq 1$. Let $\tilde{\alpha}_N[j]$, for $0 \leq j \leq N$, be continuous functions in η for some $N \geq 0$. This implies that $L_N^\eta(\tilde{\boldsymbol{\alpha}}_N)$ is also a continuous function in η. Invoking (34), we infer that $\tilde{\alpha}_{N+1}[j]$, for $j = 0, 1, \ldots, N$, are continuous functions in η.

Hence, $\sum_{j=0}^{N}\beta_N[j] \leq \frac{N+1}{Kp}$. Therefore, if $\frac{N+1}{Kp} \leq 1$ then $\sum_{j=0}^{N}\beta_N[j] \leq 1$. Hence, from Theorem 2, it follows that $\boldsymbol{\alpha}_N^* = \boldsymbol{\beta}_N$ and $\eta = 0$.

F. Derivation of (11)

Averaging the expression for the probability of success in (26) in Appendix B over the distribution of k, we get

$$P_N(\boldsymbol{\alpha}_N) = \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!}\sum_{i=0}^{N} k\alpha_N[i]\left(1-\sum_{j=0}^{i}\alpha_N[j]\right)^{k-1}. \tag{40}$$

Interchanging the two summations and using the Taylor series expansion for e^x yields (11).

G. Proof of Theorem 3

The ideas in this proof are similar to those in Appendix D. We, therefore, only highlight the key steps. For the Poisson prior, define the auxiliary function $H_N^{\gamma}(\boldsymbol{\alpha}_N)$ as

$$H_N^{\gamma}(\boldsymbol{\alpha}_N) = P_N(\boldsymbol{\alpha}_N) + \gamma e^{-\lambda\sum_{j=0}^{N}\alpha_N[j]}, \tag{41}$$

for $\gamma \geq 0$. As in the binomial prior case, we first derive the optimal $\tilde{\boldsymbol{\alpha}}_N = (\tilde{\alpha}_N[0],\ldots,\tilde{\alpha}_N[N])$ that maximizes $H_N^{\gamma}(\boldsymbol{\alpha}_N)$ and then show that it solves $\mathcal{OP}$, provided γ is chosen appropriately. The term $\gamma e^{-\lambda\sum_{j=0}^{N}\alpha_N[j]}$ can be interpreted as a displeasure function [27] that characterizes the penalty when the constraint in (13) is not met.

1) *Derivation of* $\tilde{\boldsymbol{\alpha}}_N$*:*[7] From (41) and (11), we get

$$\begin{aligned} H_N^{\gamma}(\boldsymbol{\alpha}_N) = \lambda\alpha_N[0]e^{-\lambda\alpha_N[0]} &+ \lambda\sum_{i=1}^{N}\alpha_N[i]e^{-\lambda\sum_{j=0}^{i}\alpha_N[j]} \\ &+ \gamma e^{-\lambda\sum_{j=0}^{N}\alpha_N[j]}. \end{aligned} \tag{42}$$

Taking $e^{-\lambda\alpha_N[0]}$ as a common factor from the last two terms, we get

$$\begin{aligned} H_N^{\gamma}(\boldsymbol{\alpha}_N) &= \lambda\alpha_N[0]e^{-\lambda\alpha_N[0]} \\ &\quad + e^{-\lambda\alpha_N[0]}H_{N-1}^{\gamma}(\alpha_N[1],\ldots,\alpha_N[N]), \quad (43)\\ &\leq \lambda\alpha_N[0]e^{-\lambda\alpha_N[0]} + e^{-\lambda\alpha_N[0]}H_{N-1}^{\gamma}(\tilde{\boldsymbol{\alpha}}_{N-1}), \quad (44) \end{aligned}$$

where $\tilde{\boldsymbol{\alpha}}_{N-1} = (\tilde{\alpha}_{N-1}[0],\ldots,\tilde{\alpha}_{N-1}[N-1])$ maximizes H_{N-1}^{γ}. Hence, from (43) and (44), given any $\alpha_N[0] \in (0,1)$, the upper bound in (44) can be achieved by setting

$$\alpha_N[j] = \tilde{\alpha}_{N-1}[j-1], \quad \text{for } 1 \leq j \leq N. \tag{45}$$

Furthermore, it can be shown that $\lambda\alpha_N[0]e^{-\lambda\alpha_N[0]} + e^{-\lambda\alpha_N[0]}H_{N-1}^{\gamma}(\tilde{\boldsymbol{\alpha}}_{N-1})$ is maximized when $\alpha_N[0] = \frac{1-H_{N-1}^{\gamma}(\tilde{\boldsymbol{\alpha}}_{N-1})}{\lambda}$. Thus, $\tilde{\boldsymbol{\alpha}}_N$ follows the recursion in (16). Finally, for $N=0$, $\tilde{\alpha}_0[0] = \frac{1-\gamma}{\lambda}$ maximizes $H_0^{\gamma}(\alpha_0[0]) = \lambda\alpha_0[0]e^{-\lambda\alpha_0[0]} + \gamma e^{-\lambda\alpha_0[0]}$.

For $\gamma = 0$, $H_N^{\gamma}(\boldsymbol{\alpha}_N) = P_N(\boldsymbol{\alpha}_N)$ and the recursion in (16) reduces to (15). Thus, $\tilde{\boldsymbol{\alpha}}_N = \boldsymbol{\beta}_N$, for $\gamma = 0$.

3) With the above results, we now prove the optimality of $\boldsymbol{\alpha}_N^*$. If $\sum_{j=0}^{N}\beta_N[j] \leq 1$, then $\boldsymbol{\beta}_N$ is feasible and, hence, solves $\mathcal{OP}$. Therefore, $\boldsymbol{\alpha}_N^* = \boldsymbol{\beta}_N$.

Else, if $\sum_{j=0}^{N}\beta_N[j] > 1$, then applying the intermediate value theorem [25], there exists an $\eta \in (0,1)$ such that $\sum_{j=0}^{N}\tilde{\alpha}_N[j] = 1$.[8] Clearly, $\tilde{\boldsymbol{\alpha}}_N$ is feasible. Further, for this choice of γ,

$$H_N^{\gamma}(\tilde{\boldsymbol{\alpha}}_N) = P_N(\tilde{\boldsymbol{\alpha}}_N) + \gamma e^{-\lambda}. \tag{46}$$

By definition, for any feasible $\boldsymbol{\alpha}_N$, we have $H_N^{\gamma}(\tilde{\boldsymbol{\alpha}}_N) \geq H_N^{\gamma}(\boldsymbol{\alpha}_N)$. From (41), it follows that

$$P_N(\tilde{\boldsymbol{\alpha}}_N) \geq P_N(\boldsymbol{\alpha}_N) + \gamma\left[e^{-\lambda\sum_{j=0}^{N}\alpha_N[j]} - e^{-\lambda}\right] \geq P_N(\boldsymbol{\alpha}_N), \tag{47}$$

because $\sum_{j=0}^{N}\alpha_N[j] \leq 1$ and $\gamma \geq 0$. Hence, for every $\boldsymbol{\alpha}_N$, $P_N(\tilde{\boldsymbol{\alpha}}_N) \geq P_N(\boldsymbol{\alpha}_N)$. Therefore, $\boldsymbol{\alpha}_N^* = \tilde{\boldsymbol{\alpha}}_N$.

H. Proof of Lemma 4

From (15), we have

$$\beta_N[j] = \beta_{N-1}[j-1] = \cdots = \beta_{N-j}[0], \quad \text{for } 0 \leq j \leq N. \tag{48}$$

Further, since $H_{N-1}^{\gamma}(\boldsymbol{\beta}_N) \geq 0$,

$$\beta_N[0] = \frac{1-H_{N-1}^{\gamma}(\boldsymbol{\beta}_N)}{\lambda} \leq \frac{1}{\lambda}. \tag{49}$$

Thus, $\beta_N[j] \leq \frac{1}{\lambda}$, for $0 \leq j \leq N$. Hence, $\sum_{j=0}^{N}\beta_N[j] \leq \frac{N+1}{\lambda}$. Thus, if $N+1 \leq \lambda$ then $\sum_{j=0}^{N}\beta_N[j] \leq 1$. Using Theorem 3, this implies that $\boldsymbol{\alpha}_N^* = \boldsymbol{\beta}_N$ and $\gamma = 0$.

I. Proof of Lemma 5

It suffices to prove that $\mathbb{E}_Q[S_i(f)]$ is minimized when Q is a point mass distribution, which puts the entire probability mass at one point i, for some $i \in \mathcal{K}$.

Since $\mathcal{K}$ is finite, there exists an $i^* \in \mathcal{K}$ such that $S_{i^*}(f) \leq S_i(f)$, for every $i \in \mathcal{K}$. Hence, for any distribution Q, we have $S_{i^*}(f) \leq \mathbb{E}_Q[S_i(f)]$. Further, this lower bound for $\mathbb{E}_Q[S_i(f)]$ is achieved when Q is chosen to be a point mass distribution at i^*. Hence, the result follows.

J. Proof of Theorem 4

Let $f : [0,1] \to [0,\infty)$ be an MNI mapping. Recall that $S_k(f)$ is the success probability when k nodes participate. As in Appendix A, define a new mapping g as follows:

$$g(\mu) = \begin{cases} \left\lfloor \frac{f(\mu)}{\Delta} \right\rfloor \Delta, & \text{if } f(\mu) \leq T_{\max}, \\ T_{\max} + \epsilon, & \text{if } f(\mu) > T_{\max}, \end{cases} \tag{50}$$

where $\epsilon > 0$. Using the same arguments as in Appendix A, one can show that for any $k \in \mathcal{K}$ and for any realization of the metrics, using g results in a success whenever using f results in a success. Thus, $S_k(g) \geq S_k(f)$, for all $k \in \mathcal{K}$. Hence, $\min_{k\in\mathcal{K}} S_k(g) \geq \min_{k\in\mathcal{K}} S_k(f)$. Hence, the optimal solution is an MNI staircase function.

[7] As in the binomial prior case, maximizing $H_N^{\gamma}(\boldsymbol{\alpha}_N)$ can be stated as a finite-horizon dynamic program.

[8] As in Appendix D, it can be shown that $\tilde{\alpha}_N[j]$s are continuous functions of γ. Further, $\sum_{j=0}^{N}\tilde{\alpha}_N[j] = 0$ when $\gamma = 1$.

REFERENCES

[1] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 659–672, Mar. 2006.

[2] V. Shah, N. B. Mehta, and R. Yim, "The relay selection and transmission trade-off in cooperative communication systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 8, pp. 2505–2515, Aug. 2010.

[3] D. Tse and P. Vishwanath, *Fundamentals of Wireless Communications*. Cambridge University Press, 2005.

[4] Z. Zhou, S. Zhou, J.-H. Cui, and S. Cui, "Energy-efficient cooperative communication based on power control and selective single-relay in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3066–3078, Aug. 2008.

[5] Q. Zhao and L. Tong, "Opportunistic carrier sensing for energy-efficient information retrieval in sensor networks," *EURASIP J. Wireless Commun. Netw.*, no. 2, pp. 231–241, May 2005.

[6] M. Nekovee and B. Bogason, "Reliable and effcient information dissemination in intermittently connected vehicular adhoc networks," in *Proc. 2007 IEEE VTC – Spring*, pp. 2486–2490.

[7] R. Yim, J. Guo, P. Orlik, and J. Zhang, "Received power-based prioritized rebroadcasting for V2V safety message dissemination," in *Proc. 2009 Int. Transport Syst. World Congr.*

[8] M. Khabbaz, W. Fawaz, and C. Assi, "Which vehicle to select?" *IEEE Commun. Lett.*, vol. 16, no. 6, pp. 812–815, Jun. 2012.

[9] A. Attar, N. Devroye, H. Li, and V. C. M. Leung, "A unified scheduling framework based on virtual timers for selfish-policy shared spectrum," in *Proc. 2010 ICC*, pp. 1–5.

[10] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1277–1294, Jun. 2002.

[11] V. Shah, N. B. Mehta, and R. Yim, "Optimal timer based selection schemes," *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1814–1823, Jun. 2010.

[12] R. Talak and N. B. Mehta, "Feedback overhead-aware, distributed, fast, and reliable selection," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3417–3428, Nov. 2012.

[13] C.-S. Hwang and J. Cioffi, "Opportunistic CSMA/CA for achieving multi-user diversity in wireless LAN," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2972–2982, Jun. 2009.

[14] L. Jiang and J. Walrand, "A distributed CSMA algorithm for throughput and utility maximization in wireless networks," *IEEE/ACM Trans. Netw.*, vol. 18, no. 3, pp. 960–972, Jun. 2010.

[15] ——, "Approaching throughput-optimality in distributed CSMA scheduling algorithms with collisions," *IEEE/ACM Trans. Netw.*, vol. 19, no. 3, pp. 816–829, Jun. 2010.

[16] L. Jiang, D. Shah, J. Shin, and J. Walrand, "Distributed random access algorithm: scheduling and congestion control," *IEEE Trans. Inf. Theory*, vol. 56, no. 12, pp. 6182–6207, Dec. 2010.

[17] A. Bletsas, H. Shin, and M. Win, "Cooperative communications with outage-optimal opportunistic relaying," *IEEE Trans. Wireless Commun.*, vol. 6, no. 9, pp. 3450–3460, Sep. 2007.

[18] D. P. Bertsekas and R. G. Gallager, *Data Networks*, 2nd ed. Prentice Hall, 1992.

[19] X. Qin and R. Berry, "Opportunistic splitting algorithms for wireless networks," in *Proc. 2004 INFOCOM*, pp. 1662–1672.

[20] T. Tang and R. Heath, "Opportunistic feedback for downlink multiuser diversity," *IEEE Commun. Lett.*, vol. 9, no. 10, pp. 948–950, Oct. 2005.

[21] A. J. Izenman, "Recent developments in nonparametric density estimation," *J. Am. Stat. Assoc.*, vol. 86, no. 413, pp. 205–224, Mar. 1991.

[22] V. Shah, N. B. Mehta, and R. Yim, "Splitting algorithms for fast relay selection: generalizations, analysis, and a unified view," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1525–1535, Apr. 2010.

[23] S. M. Ross, *Stochastic Processes*, 2nd ed. John Wiley and Sons, 2009.

[24] ——, *Introduction to Probability Models*, 9th ed. Academic Press, Inc., 2006.

[25] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. McGraw-Hill, 1976.

[26] R. K. Sundaram, *A First Course in Optmization Theory*. Cambridge University Press, 1996.

[27] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2003.

Rajat Talak (S'07) received his Bachelor of Technology degree in Electronics and Communication Eng. from National Institute of Technology (NITK), Karnataka, India in 2010. He is currently pursuing his Master of Science degree from the Dept. of Electrical Communication Eng., Indian Institute of Science (IISc), Bangalore, India. His current research interests include selection algorithms, cooperative communication, cellular systems, and large networks. His research inclination is towards understanding, modeling, and designing of futuristic communication systems and networks.

Neelesh B. Mehta (S'98-M'01-SM'06) received his Bachelor of Technology degree in Electronics and Communications Eng. from the Indian Institute of Technology (IIT), Madras in 1996, and his M.S. and PhD degrees in Electrical Eng. from the California Institute of Technology, Pasadena, CA, USA in 1997 and 2001, respectively. He is now an Associate Professor in the Dept. of Electrical Communication Eng., Indian Institute of Science (IISc), Bangalore, India. Prior to joining IISc, he was a research scientist in AT&T Laboratories, NJ, USA from 2001 to 2002, Broadcom Corp., NJ, USA from 2002 to 2003, and Mitsubishi Electric Research Laboratories (MERL), MA, USA from 2003 to 2007.

His research includes work on cellular systems, multiple access protocols, MIMO, energy harvesting, cognitive radio, and cooperative communications. He was also actively involved in the Radio Access Network (RAN1) standardization activities in 3GPP from 2003 to 2007. He was a TPC co-chair for tracks/symposia in COMSNETS 2014, Globecom 2013, ICC 2013, SPCOM 2012, WISARD 2010 and 2011, NCC 2011, VTC 2009 (Fall), and Chinacom 2008. He has co-authored 40+ IEEE journal papers, 70 conference papers, and three book chapters, and is a co-inventor in 20 issued US patents. He served as an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2008 to 2011, and is now an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE WIRELESS COMMUNICATIONS LETTERS, and *Journal of Communications and Networks*. He currently serves as the Director of Conference Publications on the Board of Governors of the IEEE Communications Society.