

Integrated Cost-Based MAC and Routing Techniques for Hop Count Forwarding in Wireless Sensor Networks

Michele Rossi, *Member, IEEE*, and Michele Zorzi, *Fellow, IEEE*

Abstract—This paper presents integrated MAC/routing solutions for wireless sensor networks. At the MAC layer, every node accesses the channel according to its own cost by means of properly defined cost-dependent access probabilities. Costs are used to capture the suitability of a node to act as the relay and may depend on several factors such as residual energies, link conditions, queue state, etc. Our cost-aware MAC discriminates nodes right in the channel access phase by therefore assisting the forwarding decisions to be made at the routing level. In fact, nodes with high costs are ruled out from channel contention and are not considered when making routing decisions. This provides the routing layer with better relay candidates and, at the same time, decreases the number of in-range devices contending for the channel, thereby reducing interference. The proposed MAC scheme is coupled with routing over hop count (HC) coordinates. To this end, we introduce a set of rules designed to perform HC routing by exploiting first and second order neighborhood information. These are then integrated with our MAC scheme according to a cross-layer approach and their effectiveness is demonstrated by means of analysis and simulation.

Index Terms—Communication system routing, access protocols, decision-making, cost optimal control, communication system performance.



1 INTRODUCTION

WIRELESS sensor networks (WSNs) are composed of low cost devices with limited computation and storage capabilities whose function is to gather, process, possibly aggregate, and, finally, deliver data to a central entity commonly referred to as a *sink*. Notably, WSN nodes have scarce capabilities compared to wireless ad hoc networks. However, these networks have to meet even harder requirements than their ad hoc counterparts as they are called to work unattended for long periods of time [1]. The focus of the present work is on integrated MAC and routing schemes for wireless sensor networks over hop count (HC) topologies.¹ We consider that each link is characterized by a cost which may be used to reflect transmission reliability, residual energy at receiving nodes, state of their queues, etc. As in geographical routing, we still exploit some topology-related information to select the best relay node for the current transmission. However, instead of considering node coordinates, we do so by exploiting HC values, as in [2], [3]. Our protocols are designed to work well in environments where nodes might periodically go into sleep modes for energy saving and the network topology changes due to mobility and/or node failures. In such a scenario, it makes

sense to obtain routing information on the fly (stateless schemes). For routing, we present strategies exploiting both first order (in-range nodes) and second order (two hops away nodes) information about the node costs. The latter is modeled through a statistical estimate of the cost in the second order neighborhood (two hops away). Moreover, rather than explicitly accounting for node energy and/or delay [4], [5], we deal with general cost models. In this way, we gain in generality as our approach can be applied to any user/operator specific cost function. A discussion on possible cost functions is given in Section 3. In addition, we present a MAC scheme to be integrated with the above HC routing technique. In our MAC protocol, we consider the *cost correlation* among nodes to drive the channel contention and elect a node with a small cost as the relay. The channel access scheme also exploits the new receiver-based contention paradigm introduced in [4], i.e., the active nodes within range contend to be the relay. Moreover, nodes are not equivalent, but they are ranked based on their costs and so are their access probabilities. In such a way, a preliminary routing optimization is already possible in the channel access phase. Finally, HC routing and MAC algorithms are integrated according to a cross-layer approach [6], [7] and their performance is evaluated by analysis and simulation.

The paper is organized as follows: Section 2 describes related work. In Section 3, we introduce the network and cost models that are used throughout the paper. In Section 4, we summarize hop count routing and present *online* rules to make efficient forwarding decisions over HC coordinates. In Sections 5 and 6, we propose a probabilistic MAC scheme and study its performance in Sections 6.3 and 8.1. The integrated MAC/routing algorithm is then described in

1. Hop counts are defined as the minimum number of transmissions to reach the sink from a given node.

• The authors are with the Department of Information Engineering, University of Padova, via Gradenigo 6/B-35131, Padova, Italy, and with the Consorzio Ferrara Ricerche (CFR), via Saragat 1-44100, Ferrara, Italy. E-mail: {rossi, zorzi}@dei.unipd.it.

Manuscript received 16 Dec. 2004; revised 31 Aug. 2005; accepted 18 July 2006; published online 15 Feb. 2007.

For information on obtaining reprints of this article, please send e-mail to: tmc@computer.org, and reference IEEECS Log Number TMC-0345-1204.

Section 7 and its effectiveness is demonstrated through the results reported in Section 8. Finally, in Section 9, we draw the conclusions of our work.

2 RELATED WORK

Algorithms for WSNs are constrained to be *simple*, as sensor nodes cannot afford complex operations. In addition, they must also be *efficient*, as the whole network should be able to survive unattended as long as possible. Driven by these goals, many papers [2], [3], [4], [5], [8], [9], [10], [11], [12], [13], [14] were dedicated to the study of routing schemes for such networks. One of the trends in recent research [9], [10], [4], [14], [5] is to forward packets toward the destination without using routing tables, so as to minimize the overhead associated with their creation, storage, and maintenance. These algorithms are mostly based on geographical coordinates, where every node is assumed to be aware of both its own position and the position of the *sink*; this knowledge enables online routing schemes, where the decision of the relay node can be made based on geographical advancement toward the destination (sink). In geographical routing, a sensible solution to perform data forwarding [4], [15], [16], [17] is to select the relay nodes leading to the maximum geographical progress within radio range, a technique also referred to as Most Forward within Radius [16]. Besides the maximization of the advancement, other goals, such as the maximization of node residual energies, can and should be taken into account, e.g., by periodically putting the nodes into sleep modes [4], [11]. The above mentioned routing techniques are examples of *localized routing algorithms* (LRAs), where each node can only exchange information with its first-order neighbors, i.e., the nodes within radio range. The local information exchange is, in fact, crucial to achieving *scalable* solutions which avoid the large communication overhead incurred in propagating path discovery/update messages. In [5], the authors improve the performance of LRAs by introducing the concept of *partial topology knowledge forwarding*, whose goal is to determine the optimal *knowledge range* (radio range) for each node. Their aim is to improve, where possible, the network view of a node in order to approach globally optimal routing schemes by means of “controlled” local views of the network. Zorzi and Rao [4] present a joint MAC/routing scheme exploiting geographical coordinates. In [4], the forwarding area is partitioned into a number of priority regions in order to achieve two goals: 1) to reduce the number of nodes which simultaneously transmit within a single contention round and 2) to elect a relay node providing a near maximum advancement toward the sink. In fact, priority regions are ordered according to the associated geographical advancement and the channel contention starts with the region having the highest priority (maximum advancement). Lower priority regions are considered only if higher priority regions are empty. However, geographical routing is known to be affected by the following problems: 1) Every node must be aware of its own position and of the position of the sink and 2) situations may occur where, even in a connected network, there are no nodes leading to a positive advancement toward the destination (this is often referred to as the *connectivity hole* problem [9], [10], [14]). Note that geographical coordinates are often obtained by costly

operations or by dedicated hardware (GPS) that, for certain applications, may be too expensive to install in every node or may simply not work (e.g., underwater sensor networks). As a possible solution to 1) and 2), in [2], [3], forwarding decisions are made based on *virtual* coordinates, i.e., based on topology (hops) instead of true geography. Path optimality is achieved by propagating costs so as to build a *minimum cost field*. Once such a cost field is set up, every node can communicate with the most cost efficient node within range, leading to a positive advancement (in hops) toward the destination. The main drawback of these schemes is that the cost field needs to be propagated throughout the network and periodically updated, according to the network dynamics. (Observe that this is true even if nodes are static as their on-off behavior and/or a change in the costs would change the connectivity graph.) Our present work is based on the HC concept. However, differently from [2], [3], we *do not propagate node costs* but rather take online actions which are based on *local cost information* only, thereby devising an LRA. In fact, we consider stateless schemes where forwarding decisions are made on the fly [4], [5], which is expected to outperform classical routing approaches [18], [19], especially in highly dynamic networks [20]. As far as MAC is concerned, we cite here the approaches in [21], [22], [23]. In [21], the authors propose a delay optimal CSMA-based access scheme. The objective of such an approach is to maximize the probability of electing a relay within a single contention window without, however, considering that nodes may be characterized by different capabilities and that, for this reason, may not be equivalent for forwarding purposes. Further, the authors of [22] and [23] present an IEEE 802.11-like contention algorithm where nodes contend for the channel by exploiting cost-dependent heuristics. In this paper, we present a MAC approach preserving the desirable features of previous schemes. Our objective is to elect a relay node with a low cost and, at the same time, drive the contention by means of analytically derived and cost dependent access probabilities. We complement previous work [22], [23] by dynamically adapting the channel access behavior based on the statistics governing node costs, which, in turn, is estimated online as part of the contention.

3 NETWORK MODEL

We model the network as a graph $\mathcal{G} = (\mathcal{M}, \mathcal{A})$, where \mathcal{M} is the set of nodes and \mathcal{A} is the set of “arcs” or links between nodes. Among the $m = |\mathcal{M}|$ nodes in \mathcal{M} , we consider a special device called a *sink*, with the function of gathering data. The function of the remaining $m - 1$ nodes is instead to sense the environment and forward packets to the sink. \mathcal{A} is a set of ordered pairs (i, j) , $i, j \in \mathcal{M}$. (i, j) is referred to as the link connecting node i to node j . We keep the communication/connectivity model as general as possible without making any specific assumptions on the physical channel model, modulation, and coding techniques. Therefore, any further specification on radio propagation characteristics, fading, etc. would not affect the validity of our approach. Our framework is based on the concept of *neighboring sets*, which are sets of nodes within coverage at a given time instant. We stress that neighboring sets may dynamically vary between subsequent forwarding actions,

depending on the dynamics of the connectivity graph. It would therefore be infeasible to derive these sets once for all, e.g., at the beginning of network operations, whereas it is more reasonable to obtain them *on-demand* when the forwarding decision has to be made. Each arc $(i, j) \in \mathcal{A}$ is characterized by a cost $c(i, j)$, which depends on both endpoints i and j . In particular, $c(i, j)$ may be a function of the energy needed to transmit over the link of the signal quality at node j and of the status of node j , which may, in turn, account for residual energy, congestion, and advancement toward the sink. A detailed study of how these costs can be set up and used may be found in [24], [25], [26]. As a concrete example, in [24], link costs $c(i, j)$ are calculated considering: 1) the link quality, expressed as the probability of success $P_s(i, j)$ at the receiving node j , 2) the geographical advancement associated with node j , z_j , and 3) the residual energy of node j . Advancements and link quality are jointly accounted for by means of the expected advancement $P_s(i, j)z_j$ that is subsequently coupled with the residual energy $E_{res}(j)$ at node j by means of a weighted sum $c(i, j) = \mu P_s(i, j)z_j + (1 - \mu)(E_{res}(j)/E_{init})$, where $\mu \in [0, 1]$ and E_{init} is the initial energy reserve which is assumed to be the same at every node. The same expected advancement metric is also investigated in [25], [26]. In the present paper, we keep the cost model as general as possible by not focusing on a specific cost function. In addition, with minor modifications, the general approach proposed here can be used with any of the models discussed above. For each node $i \in \mathcal{M}$, we define the set \mathcal{N}_i as containing all the neighboring nodes of node i . If $n \geq 2$ is the HC of node i , we further define $\mathcal{N}_i(n)$, $\mathcal{N}_i(n-1)$, and $\mathcal{N}_i(n+1)$ as the sets of neighbors of node i with HC are equal to n , $n-1$, and $n+1$, respectively, where $\mathcal{N}_i = \mathcal{N}_i(n-1) \cup \mathcal{N}_i(n) \cup \mathcal{N}_i(n+1)$. Moreover, a *path* of length k from node i to node d is defined as an ordered list of nodes, i.e., $\mathcal{P} = \{i, r_1, r_2, \dots, r_{k-1}, d\}$, where nodes i and d are referred to as the *source* and the *destination* node, respectively. $r_j, j \in \{1, 2, \dots, k-1\}$ are referred to as relay nodes. The cost $C(\mathcal{P})$ of path \mathcal{P} is found as²

$$C(\mathcal{P}) = c(i, r_1) + \sum_{s=1}^{k-2} c(r_s, r_{s+1}) + c(r_{k-1}, d). \quad (1)$$

The choice of selecting an additive cost function as the path cost criterion is reasonable since additive metrics arise in many settings. For example, end-to-end delay, delay jitter, total path residual energy, and reliability can all be expressed as the sum of link weights. Path reliability, which is the product of the link reliabilities, can be also reduced to an additive metric by using logarithms [27].

4 STATISTICALLY ASSISTED ROUTING

4.1 Introduction to Hop-Count-Based Routing

The sink is the initiator of the HC field distribution procedure. This algorithm is aimed at propagating a gradient, similarly to [3]. However, the gradient in our case only depends on the underlying connectivity structure

and, differently from [2], [3], [8], is just used to propagate the hop count information rather than node costs. We stress that, in general, this procedure is to be reexecuted as the network topology changes, while, in the case of static networks, it has to be run only once during the network deployment start-up phase. Throughout the paper, we assume having a consistent HC field. An example of how hop counts can be propagated and updated can be found in [2]. An investigation of the impact of hop count inconsistencies on MAC/routing performance is left for future research. Once the HC field has been set up, the key rule to route packets is to allow data forwarding toward the sink by following the HC gradient. If the HC field is properly updated, the packet will eventually reach the sink. It will be observed that the possibility of lack of connectivity is instead always present in geographical routing, where one may have to resort to using obstacle avoiding algorithms [10], [14].

4.2 Routing as a Sequential Decision Problem

We formulate the routing problem as a sequential decision problem, where, at every stage, a node has to elect the best relay node for the current packet. In particular, our focus is on *online* routing schemes, where forwarding decisions are made based on local knowledge (nodes within range) and on some statistical information about the second-order (two-hops away) neighborhood of the current node. For now, and only for theoretical purposes, we consider this information as available through an ideal MAC and at no cost. In fact, our main goal in this section is to describe a forwarding criterion able to guarantee good global solutions; how this criterion can be actually implemented by means of a practical MAC scheme is discussed in the following sections. Before proceeding with the analysis, we need to define some quantities. First of all, we assume that the currently occupied node is node $i \in \mathcal{M}$, that its hop count is $HC(i) = n$, and that the forwarding process is at stage $t \geq 0$, where time evolves one unit at every forwarding action. The problem to be solved by the decision maker is therefore to decide which is the best relay node among those in sets $\mathcal{N}_i(n)$ and $\mathcal{N}_i(n-1)$. Nodes in set $\mathcal{N}_i(n+1)$ are excluded a priori since, in normal operating conditions, they do not lead to satisfactory solutions.³ We refer to $j_{n-1}^t \in \mathcal{N}_i(n-1)$, $j_n^t \in \mathcal{N}_i(n)$, and to $c(i, j_{n-1}^t)$, $c(i, j_n^t)$ as the minimum cost nodes⁴ in sets $\mathcal{N}_i(n)$ and $\mathcal{N}_i(n-1)$ and the related costs, respectively. Formally, $j_{n-1}^t = \arg \min_{j \in \mathcal{N}_i(n-1)} \{c(i, j)\}$ and $j_n^t = \arg \min_{j \in \mathcal{N}_i(n)} \{c(i, j)\}$. We define *forwarding cycle* as the sequence of steps between the stage where a node with hop count n is reached for the first time and the instant in which a node with hop count $n-1$ is eventually selected as relay. In other words, a cycle is the number of steps (forwarding actions) it takes for the decision maker (packet) to advance one hop toward the destination. As mentioned above, the main objective of

3. This has been verified by extensive simulation and is also supported by previous studies [4].

4. In the case where there are multiple nodes with the same minimum cost in the same set, we randomly pick any of them as they are, by definition, equivalent.

2. Under the assumption of additive cost function, see [27].

the routing scheme is to minimize the total cost of the selected path (see (1)). In particular, the whole path can be decomposed as a disjoint sequence of n forwarding cycles,⁵ i.e., its cost equals the summation of the costs incurred in each forwarding cycle. Further, if we assume that the forwarding cycles are statistically independent, the minimization of the expected cost associated with the full path is obtained by minimizing the expectation of each term taken separately. In this case, devising a scheme to minimize the path cost is equivalent to devising a procedure to minimize the expected cost of each cycle. The analytical framework we developed for this purpose is discussed in Section 4.3. As the major point of this paper consists of the proposal of a MAC scheme to be coupled with hop count routing techniques, we keep the discussion in the next section at an intuitive level. A more rigorous treatment, including formal proofs, can be found in [28].

4.3 Optimal Routing Rules

Consider a generic forwarding step $t \geq 0$ and consider that the packet at time t is at node i with $\text{HC}(i) = n$ and that time 0 corresponds to the instant where the current forwarding cycle has started. At time t , the decision maker has to choose a forwarding action, i.e., whether the packet is to be forwarded to node j_{n-1}^t or j_n^t . We define the action set as $\mathcal{A}_t = \{a_{n-1}^t = j_{n-1}^t, a_n^t = j_n^t\}$ and let X_t be the decision maker's current state, which is a function of i , $\mathcal{N}_i(n)$, $\mathcal{N}_i(n-1)$, and of the costs $c(i, j)$, $j \in \{\mathcal{N}_i(n) \cup \mathcal{N}_i(n-1)\}$. Moreover, we assume that, if action $a(t) \in \mathcal{A}_t$ is chosen when in state X_t , $t \geq 0$, a cost $C(X_t, a(t)) \geq 0$ is incurred. In what follows, we discuss the first possible way to model the costs associated with this decision process. In this first approach, if $a(t) = a_n^t$, a cost $C(X_t, a_n^t) = c(i, j_n^t)$ is paid and the cycle is continued, whereas, if $a(t) = a_{n-1}^t$, the cycle ends with a final cost $C(X_t, a_{n-1}^t) = c(i, j_{n-1}^t)$. With this cost model, and considering a dynamic programming approach [29], we can write the optimality equation as follows:

$$V(X_t) = \min \left[C(X_t, a_{n-1}^t), C(X_t, a_n^t) + \int_{\mathcal{D}_X} V(X_{t+1}) dF(X_{t+1}) \right], \quad (2)$$

where $V(X_t)$ is the expected minimum cost when in state X_t , whereas the term $\int_{\mathcal{D}_X} V(X_{t+1}) dF(X_{t+1})$ accounts for the future costs incurred when the cycle is continued toward node j_n^t ($a(t) = a_n^t$), \mathcal{D}_X is the domain of X_{t+1} , and $F(X_{t+1})$ is its cumulative distribution function (cdf). Considering a one-stage look-ahead policy [29], we can manipulate (substituting $V(X_{t+1})$ with $C(X_{t+1}, a_{n-1}^{t+1})$) (2) and obtain the set

$$\begin{aligned} \mathcal{B}_1 = & \left\{ X_t : C(X_t, a_{n-1}^t) \leq C(X_t, a_n^t) + \int_{\mathcal{D}_X} C(X_{t+1}, a_{n-1}^{t+1}) dF(X_{t+1}) \right\} \\ = & \{ X_t : c(i, j_{n-1}^t) - c(i, j_n^t) \leq \mathcal{E} \}, \end{aligned} \quad (3)$$

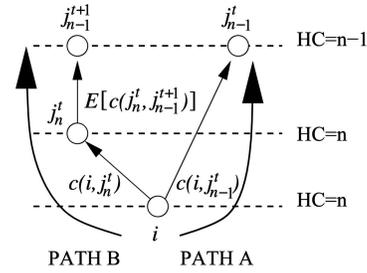


Fig. 1. Diagram for the OSP1 forwarding policy.

where \mathcal{E} is the expected value of the minimum cost at the next stage $t+1$

$$\mathcal{E} = E[c(j_n^t, j_{n-1}^{t+1})] = \int_{\mathcal{D}_c} c(j_n^t, j_{n-1}^{t+1}) dF_{\min}(c(j_n^t, j_{n-1}^{t+1})), \quad (4)$$

where $\mathcal{D}_c = [0, 1]$ as, without loss of generality, we consider normalized costs;

$$j_{n-1}^{t+1} = \arg \min_{j \in \mathcal{N}_{j_n^t}(n-1)} \{c(j_n^t, j)\}$$

and is computed by using $F_{\min}(c(j_n^t, j_{n-1}^{t+1}))$, the cdf of the cost associated with transition $j_n^t \rightarrow j_{n-1}^{t+1}$, i.e., we first select the lowest cost node in $\mathcal{N}_{j_n^t}(n)$ and then finally converge to $\text{HC } n-1$ in the next forwarding step $t+1$. Set \mathcal{B}_1 contains the states for which stopping is at least as good as continuing for one more stage and then stopping. The policy that stops the first time the process enters set \mathcal{B}_1 is called the *one-stage look-ahead policy*. According to this policy, at time t , we should select node j_{n-1}^t and end the cycle only if $c(i, j_{n-1}^t) - c(i, j_n^t) \leq \mathcal{E}$. The rationale behind this stopping rule is illustrated in Fig. 1. The starting node (decision maker) is node i and is called to minimize the expected cost incurred in advancing one HC toward the sink. In this case, the rule indicated by set \mathcal{B}_1 assesses when PATH A is preferable, in terms of cost, over PATH B. It can be shown that, in general, the one-stage look-ahead forwarding policy defined by set \mathcal{B}_1 is not globally optimal [28], but that it is optimal only when the maximum length of each forwarding cycle is 2, i.e., when, at time t , the decision maker can either choose j_{n-1}^t and end the cycle or choose j_n^t and be constrained to end the cycle (j_{n-1}^{t+1}) at stage $t+1$. This policy is referred to as the *one-stage optimal policy OSP1*. In the following, we introduce some quantities that will be used to define a more cost efficient forwarding rule. At every stage $t \geq 0$, the decision maker is asked to make a decision in set \mathcal{A}_t . If decision a_{n-1}^t is made, then node j_{n-1}^t is selected and the cycle ends with a total cost $C_{tot}(t)$:

$$C_{tot}(t) = C_{par}(t) + c(j_n^{t-1}, j_{n-1}^t), \quad (5)$$

$$C_{par}(t) = \begin{cases} 0 & t < 1 \\ \sum_{k=0}^{t-1} c(j_n^{k-1}, j_n^k) & t \geq 1, \end{cases} \quad (6)$$

where we define $j_n^{-1} = i$. If decision a_n^t is made, the cycle is continued toward node j_n^t with a partial cost $C_{par}(t+1)$. Observe that, when $C_{par}(t+1) \geq C_{tot}(t)$, there is no point in

5. When the starting node $i \in \mathcal{M}$ has $\text{HC}(i) = n$.

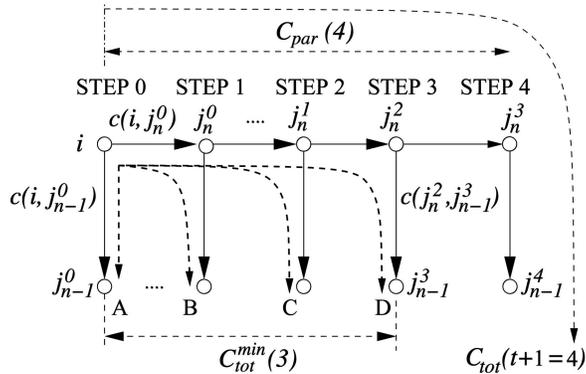


Fig. 2. Diagram for the OSP2 routing policy. The decision tree after four forwarding steps. The current stage is step $t = 3$ and $c(j_n^3, j_{n-1}^4)$ is the unknown cost of the minimum cost link leading to HC $n - 1$ at Step 4. A, B, C, and D are all the encountered but not selected paths. $C_{tot}^{min}(3)$ is evaluated as $C_{tot}^{min}(3) = \min\{\text{Cost}(A), \text{Cost}(B), \text{Cost}(C), \text{Cost}(D)\}$.

further searching for a better solution and the cycle should end. The minimum cost of the paths encountered by the decision maker up to and including time t is evaluated as $C_{tot}^{min}(t) = \min_{0 \leq k \leq t} \{C_{tot}(k)\}$. Now, in order to devise a globally optimal policy, we consider the following extended version of the forwarding problem: At the generic step t , the decision maker has to choose an action in the set $\mathcal{A}_t = (a_{n-1}^t, a_n^t)$, where a_{n-1}^t corresponds to stopping, thereby choosing node j_{n-1}^t , whereas action a_n^t means to continue the path search toward the current minimum cost node with hop count n (j_n^t). At every decision stage t , the decision maker can keep track of previous costs by evaluating the minimum cost of all paths encountered so far $C_{tot}^{min}(t)$. Now, if the optimality criterion is to *select the minimum cost path*, the corresponding one-stage look-ahead policy obeys the following stopping set [28]:

$$\mathcal{B}_2 = \left\{ X_t : C_{tot}^{min}(t) - C_{par}(t+1) \leq \mathcal{E} \right\}, \quad (7)$$

where X_t is described in terms of the costs the decision maker came across since the beginning of the current cycle, while \mathcal{E} and $C_{par}(t+1)$ are defined in (4) and (6), respectively. $C_{tot}^{min}(t)$ is the minimum cost path toward HC $n - 1$ encountered since the beginning of the cycle. The routing policy described by set \mathcal{B}_2 asserts that it is convenient to stop at time t whenever the expected cost of stopping at time $t + 1$, $C_{par}(t+1) + \mathcal{E}$, is greater than or equal to the minimum cost of all paths encountered at times 0 through t , $C_{tot}^{min}(t)$. The aim of this rule is to continue the cycle until the expected cost of the path in the next step is higher than the minimum path cost measured so far. Observe that this rule is a refinement of OSP1. In fact, in addition to future costs, the cost history from the beginning of the current cycle (starting at time 0) is also taken into account. For cycles longer than two stages, the refined policy outperforms OSP1: Intuitively, this is due to the fact that, at the generic time t , when costs $c(i, j_n^t)$ and \mathcal{E} are very low, OSP1 continues the forwarding cycle toward node j_n^t , regardless of the accumulated cost up to time t . However, the accumulated costs may exceed a past offer (one of the paths encountered but not selected between time 0 and the

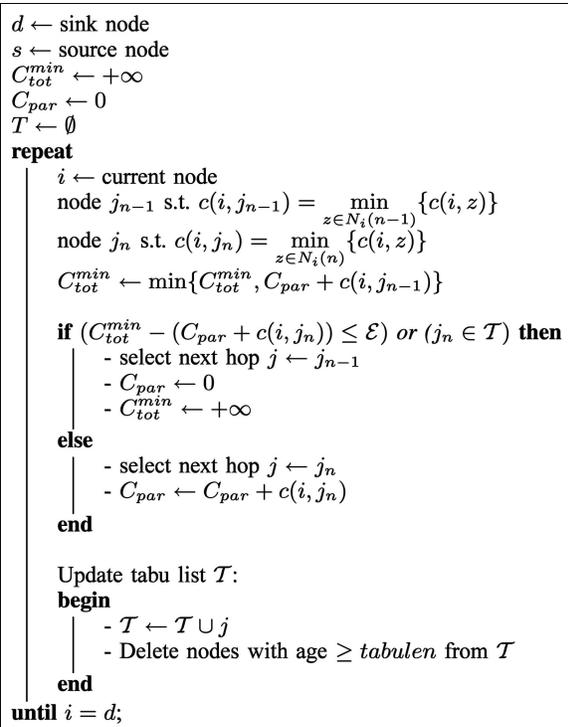


Fig. 3. Algorithm 1: Statistically-Assisted Routing Algorithm (SARA). We account for a tabu list T to avoid loops or ping-ponging between nodes

current step t) and, in this case, the best decision is to end the cycle. It can be proven [28] that the policy dictated by set \mathcal{B}_2 is globally optimal. This optimality means that the policy gives the lowest long-term expected cost [29] among all online policies exploiting the cost information of nodes in range and two-hops away, under the fundamental assumptions that nodes in $\mathcal{N}_i(n+1)$ are not considered and that the average $E[c(j_n^t, j_{n-1}^{t+1})]$ is only a function of the hop count, n . We refer to the policy described by set \mathcal{B}_2 as OSP2: Refer to Fig. 2 for a graphical representation of the decision tree and related quantities after four forwarding steps. The routing schemes defined by OSP1 (\mathcal{B}_1) and OSP2 (\mathcal{B}_2) are named Statistically Assisted Routing Algorithms (SARA) as they exploit two-hop neighbors' statistical measures (\mathcal{E}) to carry out the relay election. The full version of the SARA routing algorithm exploiting the stopping rule of (7) is reported in Algorithm 1 in Fig. 3. Finally, we observe that, differently from the cost model that we consider in this paper, where we assume independence among subsequent forwarding steps, real situations may show spatial correlation. That is, the cost statistics at future stages may depend on the current routing choice. However, as our algorithms are designed to go around high cost nodes, we expect them to still work satisfactorily also in the presence of highly correlated regions. Assume, for example, that nodes in $\mathcal{N}_i(n-1)$ have high costs. Then, due to the high correlation, we may expect their descendants to have high costs as well and, in such a case, selecting a node in $\mathcal{N}_i(n)$ with smaller cost would increase the probability of circumnavigating the high cost region, thereby leading to better performance. We leave a detailed investigation of this matter for future research.

5 A DISTRIBUTED CHANNEL ACCESS TECHNIQUE

Next, we propose and validate a MAC scheme that we will couple with the routing techniques discussed above. We want to devise a simple channel access procedure in order to discover the minimum cost nodes in sets $\mathcal{N}_i(n)$ and $\mathcal{N}_i(n-1)$. This procedure should be distributed, i.e., the decision on whether a node has to access the channel should be made independently at every node and exploiting a minimal amount of information regarding the state of neighboring devices. Also, the algorithm should be efficient in the sense that 1) the number of contentions should be small and 2) the cost of the elected relay should be reasonably close to the minimum in its set. Hence, to pick the best relay, a given node i needs to first locate the minimum cost nodes in the two sets $\mathcal{N}_i(n)$ and $\mathcal{N}_i(n-1)$. Subsequently, either (3) or (7) can be used at node i to make forwarding decisions. The problem to be solved is therefore to provide an efficient way to discover the minimum cost nodes in the two above sets. A possible integrated MAC/routing is outlined next.

1. Node i sends a *request* message (REQ) addressed to all nodes in $\mathcal{N}_i(n-1)$. The node will include its HC (n) in the REQ and an HC identifier addressing all devices in $\mathcal{N}_i(n-1)$. Estimates for the number of nodes (N) in $\mathcal{N}_i(n-1)$ and correlation among costs (ρ) are also included in the REQ.
2. Nodes in $\mathcal{N}_i(n-1)$ contend for the channel. Every node $j \in \mathcal{N}_i(n-1)$ decides to reply to the REQ using a channel access probability function $P_a^j = P_a(c(i,j), N, \rho)$ which is a function of the link cost, N and ρ . This is a receiver-based contention in line with [4].
3. If multiple nodes reply to the REQ, a collision may occur. In case of collision, node i sends a further REQ by modifying the ρ estimate until a winner can be elected.
4. Steps 1 through 3 are repeated for set $\mathcal{N}_i(n)$. The differences in this case are that the replies should also contain the quantity \mathcal{E} , which can be estimated at every node. Also, the REQ will be processed by all nodes in $\mathcal{N}_i(n)$: Node i will include its HC and an HC identifier addressing all nodes in $\mathcal{N}_i(n)$. Finally, either (3) or (7) is used to select the relay node.

The above procedure is further detailed in Section 7. In the following section, we propose a set of functions to compute the access probability $P_a(\cdot)$ by qualitatively and quantitatively showing the role of ρ and N and highlighting the differences with respect to previous work [21].

6 COMPUTATION OF CHANNEL ACCESS PROBABILITIES

To implement our routing rules in an integrated MAC/routing scheme, we need a simple and effective procedure to discover the minimum cost nodes in the above defined forwarding sets $\mathcal{N}_i(n)$ and $\mathcal{N}_i(n-1)$, given that i is the current node. Observe that, as this node never changes during the operation, we can safely omit the dependence on node i by referring to the cost of the minimum cost

node j in a given set as c_j . In order to illustrate the problems involved in our task, let us consider a generic set of N nodes \mathcal{S}_N and consider that all nodes reply with the same channel access probability $P_a^j = P_a, \forall j \in \mathcal{S}_N$. Moreover, let us define the success probability P_{succ} as the probability that a single node replies to the REQ, whereas all the other nodes stay silent, each with probability $(1 - P_a)$: $P_{succ} = NP_a(1 - P_a)^{N-1}$. In this case, it is easy to verify [21] that P_{succ} is maximized as $P_a = 1/N$ and its maximum value is given by $P_{succ}^{max} = (1 - 1/N)^{N-1}$. However, we shall observe that this methodology does not provide a way of discriminating the minimum cost node in \mathcal{S}_N . In fact, all nodes access the channel with the same probability, irrespective of their costs. Given that, it is reasonable to think of a distributed scheme where such probabilities differ and are computed depending on node costs, i.e., if $c_j \in [0, 1]$ is the cost of node $j \in \mathcal{S}_N$, we could think of a cost dependent channel access function $P_a^j = P_a(c_j)$. Next, we devise appropriate functions and techniques to relate node costs to channel access probabilities so as to keep P_{succ} at a reasonably high value while ensuring that the cost of the winner is sufficiently close to the minimum in \mathcal{S}_N . In Section 6.1, we treat the case of independent costs. The case of correlated costs is investigated in Section 6.2.

6.1 Case I: Independent and Uniformly Distributed Costs

We consider here the simplest but still important case where node costs are independent of each other and uniformly distributed. Consider the generic set \mathcal{S}_N of N nodes within coverage and consider that the cost of the generic node $k \in \mathcal{S}_N$, c_k , is distributed according to a uniform distribution in $[0, 1]$. Owing to the independent cost assumption, we can evaluate the probability $P_{min}^k(c_k)$ of node k being the minimum cost node in \mathcal{S}_N as

$$P_{min}^k(c_k) = (1 - F_C(c_k))^{N-1} = (1 - c_k)^{N-1}, \quad c_k \in [0, 1], \quad (8)$$

where $F_C(c) = \text{Prob}\{C \leq c\}$. Equation (8) corresponds to the probability that the remaining $N - 1$ nodes in \mathcal{S}_N have a cost higher than or equal to c_k . Observe that it could be reasonable to use $P_{min}^k(c_k)$ as the channel access probability P_a^k at node k . To clarify this assertion, let us define the MAC success probability P_{succ} as the probability that a single node in set \mathcal{S}_N accesses the channel. If we choose $P_a^k = P_{min}^k(c_k)$, the expected value of this probability is

$$\begin{aligned} \bar{P}_{succ} &= \underbrace{\int_0^1 \dots \int_0^1}_{N \text{ times}} \sum_{k=1}^N P_a^k \prod_{j \neq k} (1 - P_a^j) \, dc_1 \dots dc_N \\ &= \left(1 - \frac{1}{N}\right)^{N-1}. \end{aligned} \quad (9)$$

Note that, in case costs are independent and uniformly distributed, setting $P_a^k = P_{min}^k(c_k)$ (see (8)) leads to an average success probability \bar{P}_{succ} equal to the P_{succ} obtained when all nodes access the channel with the same $P_a = 1/N$ (see Section 6 above).

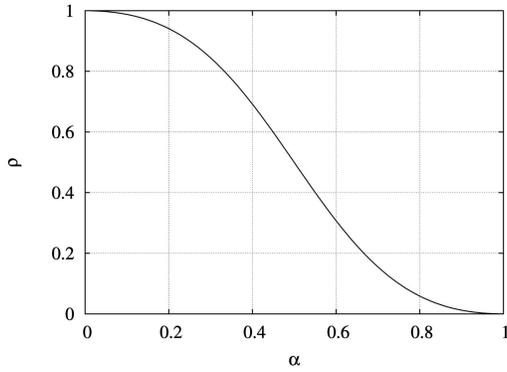


Fig. 4. Cost correlation function ρ as a function of α .

6.2 Case II: Correlated Costs

Next, we introduce a simple analytical model to take into account the cost correlation ρ between nodes within coverage and evaluate its impact on the MAC performance. As above, we consider a generic set \mathcal{S}_N of N nodes, where we refer to c_k as the cost associated with node $k \in \mathcal{S}_N$. Moreover, we assume that c_k is given by $c_k = \bar{c} + \gamma_k$, where $\bar{c} \in [0, 1]$ is a cost component common to all nodes in \mathcal{S}_N , whereas γ_k is an additive random displacement uniformly distributed in $[-\alpha\bar{c}, \alpha(1 - \bar{c})]$, $\alpha \in [0, 1]$, and independently picked for every node k . For instance, $\alpha = 0$ corresponds to the fully correlated case as all node costs collapse to \bar{c} . On the other hand, $\alpha = 1$ gives the i.i.d. case ($\rho = 0$) as the costs of every pair of nodes in \mathcal{S}_N are uncorrelated. Intermediate values of α lead to a correlation $\rho \in (0, 1)$. The correlation coefficient between any two nodes $r, s \in \mathcal{S}_N$ is $\rho_{r,s} = (E[c_r c_s] - E[c_r]E[c_s]) / (\sigma_r \sigma_s)$ where $\sigma_s^2 = E[(c_s - E[c_s])^2]$. It can be verified that (Fig. 4)

$$\rho_{r,s} = \frac{(1 - \alpha)^2}{(1 - \alpha)^2 + \alpha^2}. \quad (10)$$

The parameter α can be easily mapped as a function of $\rho = \rho_{r,s}$ as follows:

$$\alpha = \begin{cases} \frac{\rho - 1 + \sqrt{\rho(1 - \rho)}}{2\rho - 1} & 0 \leq \rho < 1/2 \text{ and } 1/2 < \rho \leq 1 \\ 1/2 & \rho = 1/2. \end{cases} \quad (11)$$

In the following, we consider a generic node $k \in \mathcal{S}_N$ with cost c_k and we compute the probability that node k is the minimum cost node in \mathcal{S}_N , $P_{min}^k(c_k)$. According to the above cost model $P_{min}^k(c_k)$ can be found in close form. For the sake of clarity, we give here the final result, while the reader is referred to the Appendix for its derivation.

$$P_{min}^k(c_k) \Big|_{0 < \alpha \leq 0.5} = \begin{cases} \frac{\alpha^N - (\alpha - c_k)^N}{c_k N \alpha^{N-1}} & 0 \leq c_k < \alpha \\ N^{-1} & \alpha \leq c_k \leq 1 - \alpha \\ \frac{(1 - c_k)^{N-1}}{N \alpha^{N-1}} & 1 - \alpha < c_k \leq 1, \end{cases} \quad (12)$$

$$P_{min}^k(c_k) \Big|_{0.5 < \alpha < 1} = \begin{cases} \frac{\alpha^N - (\alpha - c_k)^N}{c_k N \alpha^{N-1}} & 0 \leq c_k < 1 - \alpha \\ \frac{(1 - c_k)^N - (\alpha - c_k)^N}{(1 - \alpha) N \alpha^{N-1}} & 1 - \alpha \leq c_k \leq \alpha \\ \frac{(1 - c_k)^{N-1}}{N \alpha^{N-1}} & \alpha < c_k \leq 1. \end{cases} \quad (13)$$

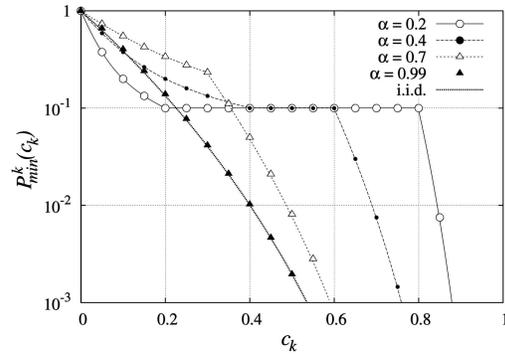


Fig. 5. $P_{min}^k(c_k)$ as a function of c_k for $N = 10$ and $\alpha \in \{0.2, 0.4, 0.7, 0.99\}$. The independent cost case is also plotted for comparison.

For the limiting cases $\alpha = 0$ and $\alpha = 1$, $P_{min}^k(c_k)$ reduces to

$$P_{min}^k(c_k) \Big|_{\alpha=0} = \frac{1}{N}, \quad P_{min}^k(c_k) \Big|_{\alpha=1} = (1 - c_k)^{N-1}. \quad (14)$$

Observe that $P_{min}^k(c_k)$ is continuous in c_k and that, in (12) and (13), $\lim_{c_k \rightarrow 0} P_{min}^k(c_k) = 1$ and $P_{min}^k(1) = 0$, i.e., a node with $c_k = 0$ is, with probability one, among the minimum cost nodes in \mathcal{S}_N , whereas, with $c_k = 1$, the probability of being the minimum cost node is zero (when $\alpha > 0$). Equations (12) and (13) are therefore consistent with the notion of minimum cost node.

Fig. 5 reports the above probability $P_{min}^k(c_k)$ as a function of c_k for some significant values of α . The curve representing the independent case (i.i.d., i.e., $\rho = 0$ or $\alpha = 1$) is drawn for comparison. Note that, as ρ tends to 1 ($\alpha \rightarrow 0$), $P_{min}^k(c_k)$ approaches $1/N$, i.e., the optimal access probability in the fully correlated case. With a decreasing ρ , instead, the curve still assumes the value $1/N$ in the interval $c_k \in [\alpha, 1 - \alpha]$ (see (12)), whereas the probability of being the minimum cost node quickly goes to zero as c_k increases beyond $(1 - \alpha)$ and tends to 1 as $c_k \rightarrow 0$. This behavior continues for a decreasing ρ (increasing α), up to the breaking point $\rho = \alpha = 0.5$, where the portion of the curve with value $1/N$ collapses to a single point. Further, as ρ continues to decrease, the behavior of the curve changes according to (13), getting closer to the fully independent ($\rho = 0$) situation. Fig. 5 highlights the substantial differences in $P_{min}^k(c_k)$ as a function of ρ . Next, we show how these functions can be exploited to improve the channel access phase.

6.3 Performance Comparison of the Proposed Channel Access Methods

In this section, we derive an exact expression for the success probability P_{succ} as a function of the cost correlation ρ and of the selected channel access probability function $P_a(\cdot)$. Besides, we also characterize the goodness of the solution found by checking how close we can get to the actual minimum cost. With the correlation model introduced in Section 6.2, the average successful access probability can be written as

$$P_1 = \int_0^1 \underbrace{\int_{\bar{c} - \alpha \bar{c}}^{\bar{c} + \alpha(1 - \bar{c})} \cdots \int_{\bar{c} - \alpha \bar{c}}^{\bar{c} + \alpha(1 - \bar{c})}}_{N \text{ times}} P_{succ}(c_1, \dots, c_N, \alpha, N) \alpha^{-N} \mathbf{d}c_1 \cdots \mathbf{d}c_N \mathbf{d}\bar{c}, \quad (15)$$

where the success probability function $P_{succ}(c_1, \dots, c_N, \alpha, N)$ is averaged over all possible values of the node costs, whose joint pdf (conditioned on \bar{c}) is given by α^{-N} (see (20)).

$$P_{succ}(c_1, \dots, c_N, \alpha, N) = \sum_{k=1}^N \left\{ P_a^k(c_k, \alpha, N) \prod_{j \neq k} [1 - P_a^j(c_j, \alpha, N)] \right\}. \quad (16)$$

Owing to the fact that, for any given \bar{c} , the additive random displacements γ s are independent and, considering the same functional form for the access probability to be used at nodes k and j (P_a only depends on the node cost and not on the node index), the integral is simplified as

$$P_1 = \frac{N}{\alpha^N} \int_0^1 \int_{\bar{c}-\alpha\bar{c}}^{\bar{c}+\alpha(1-\bar{c})} P_a(x, \alpha, N) \mathbf{d}x \left[\int_{\bar{c}-\alpha\bar{c}}^{\bar{c}+\alpha(1-\bar{c})} (1 - P_a(y, \alpha, N)) \mathbf{d}y \right]^{N-1} \mathbf{d}\bar{c}. \quad (17)$$

P_1 is the average success probability, i.e., the probability of having a single node replying to the REQ in a given contention round. Observe that, once the node cost statistics are given, P_1 depends on the specific shape of $P_a(\cdot)$, which is used at every terminal, to decide whether to participate in the contention. For $P_a(\cdot)$, we consider the following three important cases:

1. IID: Every user replies to a REQ message with probability given by $P_a(\cdot)$ in (8). In this case, each node considers node costs to be i.i.d. The number of nodes N and its own cost are the only quantities required to determine the channel access probability.
2. ACE: The acronym stands for *adaptive correlation estimation* as, in this scheme, the nodes reply to the REQ with probability given by P_a in (12), (13), and (14), where ρ is an estimate of the actual cost correlation. As will be shown in the next sections, it is sensible to adaptively modify such a correlation estimate as the channel contention evolves. N and node cost are also exploited to calculate P_a .
3. 1/N: Each node replies to the REQ with a probability $P_a = 1/N$. As discussed above, this strategy corresponds to the optimal behavior when node costs are not relevant [21], i.e., when the objective is to maximize P_{succ} irrespective of the cost associated with the winner of the contention.

In Fig. 6, we report the success probability P_1 (i.e., exactly one node replies to the REQ) for the three aforementioned cases for $P_a(\cdot)$. As expected, we observe that the case where $P_a = 1/N$ is an upper bound for P_1 . We also note that cost correlation strongly affects the accuracy of the IID solution, whereas ACE stays quite close to the optimal scheme ($P_a = 1/N$) for any ρ . Moreover, we observe that ACE presents a minimum in $\rho = 0.5$. This minimum is due to the following facts: As $\rho < 0.5$ ($\alpha > 0.5$, (13)), the shape of $P_a(\cdot)$ is similar to the one in (8), which holds for independent costs. However, as ρ increases beyond 0.5 (see (12)), we observe the presence of a flat region ($P_a = 1/N$), as for the fully correlated cost case. Hence, for $\rho > 0.5$, the single node access behavior in the

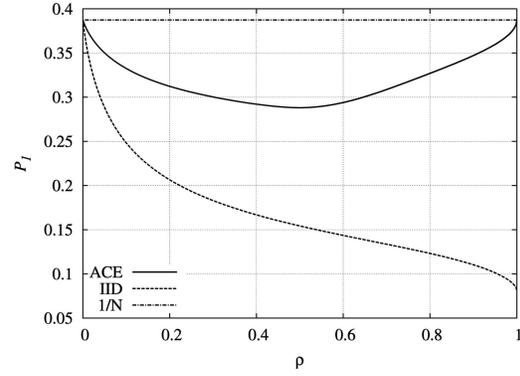


Fig. 6. P_1 as a function of ρ for $N = 10$.

ACE case changes sharply, giving rise to the slope change observed in Fig. 6. However, besides the probability of a successful channel access, a crucial part of the routing scheme is to promote a winner whose cost is sufficiently close to the minimum in the forwarding set. We therefore define a further metric which is related to the probability of providing good solutions in terms of cost. To this end, we define the following two events:

- $S = \{\text{the contention is successful}\}$. This event corresponds to the case where a single node replies to the REQ and all other nodes stay silent.
- $M = \{c_{winner} \leq c_{min} + e\}$, where $e \geq 0$ and represents our optimality criterion, while c_{winner} and c_{min} are the cost of the winner of the contention and the cost of the minimum cost node, respectively. This event is introduced to get an indication of how close c_{winner} is to c_{min} . If $e = 0$, M is verified if the winner of the contention corresponds to the minimum cost node; otherwise, if $e > 0$, we relax our optimality requirements.

The next step is to find the following probabilities: $P_2 = \text{Prob}\{M|S\}$ and $P_3 = \text{Prob}\{M \text{ and } S\}$. P_3 is derived by slightly modifying (17) as follows:

$$P_3 = \frac{N}{\alpha^N} \int_0^1 \int_{\bar{c}-\alpha\bar{c}}^{\bar{c}+\alpha(1-\bar{c})} P_a(x, \alpha, N) \left[\int_{\max(x-\bar{c}, \bar{c}-\alpha\bar{c})}^{\bar{c}+\alpha(1-\bar{c})} (1 - P_a(y, \alpha, N)) \mathbf{d}y \right]^{N-1} \mathbf{d}x \mathbf{d}\bar{c}. \quad (18)$$

P_2 is found via Bayes' formula as $P_2 = P_3/P_1$, where $P_1 = \text{Prob}\{S\}$. In Figs. 7 and 8, P_2 and P_3 are plotted as a function of ρ for the IID, ACE, and 1/N cases. $e = 0$ in Fig. 7 corresponds to the probability that the winner is also the minimum cost node, conditioned on S . We stress that the optimal case for P_1 (1/N) does poorly in terms of the goodness of the solution found. This is reasonable as, when $P_a = 1/N$ node costs are irrelevant in the channel access phase. On the other hand, the best scheme for P_2 is almost always IID, where node costs are assumed to be i.i.d. This fact is also reasonable since the $P_a(\cdot)$ curve, in this case, is the one that most penalizes high cost nodes (see Fig. 5). Furthermore, considering that $e = 0.05$ is equivalent to admitting that a winner with cost $c_{winner} \leq c_{min} + 0.05$ is still a satisfactory solution. In this case, P_2 increases as well as the gap between 1/N and ACE at low ρ values. In any

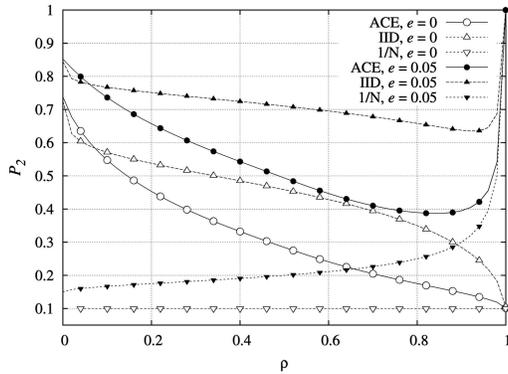


Fig. 7. $P_2 = \text{Prob}\{M|S\}$ as a function of ρ for $N = 10$.

event, this figure tells us that scheme $1/N$ should not be considered when our objective is to achieve a good solution in terms of cost. Also, one may think that IID gives the best performance in terms of the overall quality of the solution found; however, this is not true since our task here is to jointly maximize the probability of having a successful contention round and that the winner of such a contention is a satisfactory solution (S and M are jointly verified). As an example, we plot this last probability (P_3) in Fig. 8 for $e = 0.05$, where it is clear that ACE is the best scheme when the probability of success and goodness of the solution found are considered jointly. The gap is very significant between ACE and $1/N$ at $\rho = 0$ and between ACE and IID for fully correlated costs. For intermediate values of ρ , gains are lower but still important. As a last observation, it is worth noting that what is reported in the figure is obtained from statistical averaging and does not reveal what happens for limiting and specific cases. For instance, as $\rho = 1$ and \bar{c} is close to one, the IID scheme would lead to an extremely low success probability, whereas ACE in these settings degenerates to $1/N$, thereby leading to optimal performance. The gap between IID and ACE, in this case, would be even larger than in Fig. 8 and, clearly, the IID behavior in such a case should be avoided since it is highly inefficient.

7 INTEGRATED HOP COUNT MAC/ROUTING SCHEMES

In this section, we propose two integrated MAC/routing schemes. The first scheme, called SARA1, exploits the OSP2 rule in (7) and is presented below:

1. Assume that the forwarding process is currently at node i with $\text{HC}(i) = n$ and that the quantities C_{tot}^{min} and C_{par} have been computed by node i as illustrated in Algorithm 1 in Fig. 3.
2. Node i initiates the contention by sending a $\text{REQ}(n, n-1, \rho=0, N, \mathcal{T})$ to trigger a reply from all active nodes in set $\mathcal{N}_i(n-1)$. N is an estimate of the number of nodes in $\mathcal{N}_i(n-1)$ and ρ is the estimated cost correlation (initially set to 0). The REQ also includes a tabu list \mathcal{T} with the identifiers of the last $tabulen$ visited nodes (see Algorithm 1 in Fig. 3). Note that node i must also include in the REQ the information needed for the calculation of the link cost $c(i, j)$ at the receiving node j .

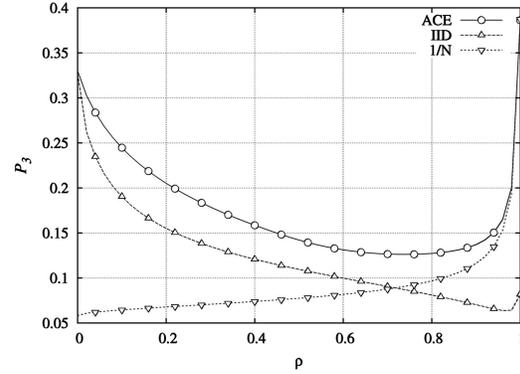


Fig. 8. $P_3 = \text{Prob}\{M \text{ and } S\}$ as a function of ρ for $N = 10$.

3. Every active node $j \in \mathcal{N}_i(n-1)$ computes $P_a = P_a(c(i, j), \rho, N)$, where $c(i, j)$ is the cost of link (i, j) and ρ and N are the estimates contained in the REQ. P_a is calculated according to (12), (13), and (14) using $c(i, j)$ instead of c_k . All nodes $j \in \mathcal{N}_i(n-1)$ reply to the REQ with probability P_a . If $j \in \mathcal{T}$, P_a is forced to be zero. In its *reply* (REP), every node includes its own identifier and cost.
4. The following three events can occur:
 - *collision*: more than one node in $\mathcal{N}_i(n-1)$ reply to the REQ and no reply can be received;
 - *silence*: no nodes reply;
 - *success*: only one node, say node j_{n-1}^* , replies to the REQ.⁶

If either *collision* or *silence* occurs, the sender updates $\rho \leftarrow \min(1, \rho + \Delta\rho)$, sends a REQ including the new ρ , and the contention is continued from Step 2. In case of *success*, node j_{n-1}^* is the winner of the contention and, at node i , $c(i, j_{n-1}) \leftarrow c(i, j_{n-1}^*)$. This last case ends the contention.

5. Node i initiates a second contention phase by sending a $\text{REQ}(n, n, \rho=0, N, \mathcal{T})$ to trigger a reply from all active nodes in set $\mathcal{N}_i(n)$. N is an estimate of the number of nodes in $\mathcal{N}_i(n)$. This contention phase proceeds as the previous one (Steps 2-4) with the only difference being that, in this case, every node includes the quantity \mathcal{E} (see (4)) in its REP. If the winner of this contention is node j_n^* , then $c(i, j_n) \leftarrow c(i, j_n^*)$.
6. C_{tot}^{min} is updated as

$$C_{tot}^{min} \leftarrow \min\{C_{tot}^{min}, C_{par} + c(i, j_{n-1}^*)\}.$$

The relay node is chosen according to (7), i.e., node j_{n-1}^* is selected if $C_{tot}^{min} - (C_{par} + c(i, j_n^*)) \leq \mathcal{E}$; otherwise, j_n^* is selected.

The second scheme, called SARA2, exploits the OSP1 rule in (3) and is detailed next.

1. Assume that the forwarding process is currently at node i with $\text{HC}(i) = n$.
2. Node i initiates the contention by sending a $\text{REQ}(n, \rho=0, N, \mathcal{T})$ to trigger a reply from all active

6. The capture effect, which may allow correct reception even in the presence of a collision, is not considered here for ease of explanation, but can be easily included.

nodes in sets $\mathcal{N}_i(n-1)$ and $\mathcal{N}_i(n)$. N is an estimate of the number of nodes in $\mathcal{N}_i(n-1) \cup \mathcal{N}_i(n)$ and ρ is the estimated cost correlation (initially set to 0). As above, the REQ also includes a tabu list \mathcal{T} .

3. Every active node $j_1 \in \mathcal{N}_i(n-1)$ computes $P_a = P_a(c(i, j_1), \rho, N)$, where $c(i, j_1)$ is the cost of link (i, j_1) , and ρ and N are the estimates contained in the REQ. P_a is calculated according to (12), (13), and (14) using $c(i, j_1)$ instead of c_k . Every active node $j_2 \in \mathcal{N}_i(n)$ computes $P_a = P_a(c'(i, j_2), \rho, N)$, where $c'(i, j_2) = \min(c(i, j_2) + \mathcal{E}, 1)$, $c(i, j_2)$ is the cost of link (i, j_2) and \mathcal{E} is as specified in (4). If $j_2 \in \mathcal{T}$, P_a is forced to be zero. All nodes in $\mathcal{N}_i(n-1)$ and $\mathcal{N}_i(n)$ reply to the REQ with their respective probability P_a .
4. As above, the following three events can occur: *collision*, *silence*, and *success*. If either *collision* or *silence* occurs, the sender updates $\rho \leftarrow \min(1, \rho + \Delta\rho)$, and a new contention round is initiated by sending a further REQ which includes the new ρ . The *success* case ends the contention and the winning node is elected as the next hop.

The following observations are in order. Starting the contention phase with $\rho = 0$ is equivalent to assuming i.i.d. costs. This makes sense as $\rho = 0$ corresponds to the channel access probability curve maximizing $P_2 = \text{Prob}\{M|S\}$ (Fig. 7). Therefore, we start the contention phase by privileging the goodness of the solution found and, subsequently, in case of *collision* or *silence*, we modify ρ to increase the *success* probability. Moreover, we note that, in the above scheme, ρ is the only parameter that is adaptively modified, whereas N is kept constant. The main reason for this is due to the sensitivity of $P_a(\cdot)$ against N , which is considerably lower than its sensitivity against ρ .⁷ Note that N and \mathcal{E} can be estimated at every node. Sender initiated algorithms for the estimation of N can be found in [30]. For \mathcal{E} , one can low pass filter the minimum costs measured at every forwarding step. Throughout the paper we will assume to have perfect estimates for both \mathcal{E} and N . The issues related to their estimation are left for future research. We observe that, in the above procedure, routing information is gathered every time a packet is to be forwarded. This is in line with recent forwarding paradigms for WSNs [4], [25], [26], [31]. In the presence of slowly varying environments, it may still make sense to retain some state as the optimality of some decisions may hold across multiple packets. In this paper, we focus on stateless solutions for highly dynamic environments and leave consideration of this trade-off for future study. Also, the OSP1 rule may occasionally lead to long sequences of nodes in $\mathcal{N}_i(n)$ [28]. To avoid this, we implemented the following modification to the basic SARA2 algorithm presented above: We additionally keep track of the number of consecutive forwarding actions to a node with the same HC level n . This information is included into REQ messages. As a packet is forwarded for a maximum number of times f_{max} to nodes belonging to the same HC level n , the next hop is

chosen among the nodes in $\mathcal{N}_i(n-1)$. In this case, the contention is carried out by sending a special REQ addressing the nodes in $\mathcal{N}_i(n-1)$ only (N in this case is an estimate for the number of nodes in this set). The version of the scheme evaluated in Section 8 includes this modification with $f_{max} = 2$. Finally, we observe that interference between different senders accessing the channel at the same time is not accounted for by the above procedure. However, this interference can be mitigated using standard techniques [31] which can be easily applied to our MAC.

8 RESULTS

As a reference model for the performance evaluation, we consider a random topology network, where nodes are uniformly deployed with normalized node density $\lambda_n = \lambda\pi R^2$, where λ is the average number of nodes per unit area, whereas R is the transmission range, assumed here to be constant. We consider a unit disk connectivity model [32], i.e., two nodes can communicate iff their distance is lower than or equal to R . Observe that the schemes presented in this paper can work for any topology setting as a different choice of the node density λ and of the connectivity model just translates into different neighboring sets. Hence, λ_n can be seen as the average number of nodes actually awake (active) within coverage. Here, we consider $\lambda_n = 15$, $R = 1$, all nodes are assumed to be randomly placed in a square area of $18R \times 18R$ and to have a perfect estimate of \mathcal{E} . In Section 8.1, we test ACE-MAC considering a single contention round. In Section 8.2, we focus on the performance of SARA1 and SARA2.

8.1 ACE-MAC: Single Contention Performance

We consider a set \mathcal{S}_N of N nodes. As previously assumed, we refer to the cost of node $j \in \mathcal{S}_N$ as c_j by omitting the index of the source node, which remains constant over a single contention phase. In the following, we assume that the sender has a perfect knowledge of N , whereas the correlation estimate sent along with REQ messages is computed as in Section 7. For the node costs, we consider the analytical model (AN) presented in Section 6.2. In addition, in order to test the goodness of our approach in a more general setting and to give evidence that it is not exclusively tailored to the simple model of Section 6.2, we also consider the NORTA [33], [34] framework, which is a known method for representing random vectors whose component random variables have arbitrary marginal distributions and correlation matrix. We define \mathbf{X} as an N -sized vector of correlated r.v. (to track the costs of nodes in \mathcal{S}_N) and $\Sigma_{\mathbf{X}}$ as its correlation matrix. Moreover, we define \mathbf{Z} as a further vector of Gaussian random variables with correlation matrix $\Sigma_{\mathbf{Z}}$. The NORTA procedure can be summarized as follows:

1. Determine a lower-triangular, nonsingular factorization \mathbf{M} of $\Sigma_{\mathbf{Z}}$ so that $\mathbf{M}\mathbf{M}^T = \Sigma_{\mathbf{Z}}$.
2. Generate $\mathbf{W} = (W_1, W_2, \dots, W_N)^T$, an $N \times 1$ vector whose elements are sampled from i.i.d. standard normal random variables and set $\mathbf{Z} \leftarrow \mathbf{M}\mathbf{W}$.

⁷ We have verified this through simulation. The detailed results are not reported here due to space constraints. However, an idea of the sensitivity against N and ρ can be gained from Fig. 5, where the dependence on N is mainly given by the horizontal line $1/N$.

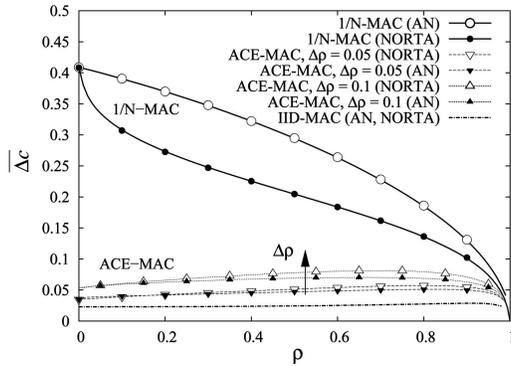


Fig. 9. Quality of the solution found within a single contention phase. Number of nodes $N = 10$, $\Delta\rho \in \{0.05, 0.1, 0.2\}$.

- Return $\mathbf{X} = (X_1, X_2, \dots, X_N)^T$, where $X_i \leftarrow F_{X_i}^{-1}[\Phi(Z_i)]$, $F_{X_i}[\cdot]$, and $\Phi(\cdot)$ are the cdfs associated with X_i and Z_i , respectively.

More details on how Σ_X can be mapped into Σ_Z can be found in [33], [34]. Here, we consider that X_i s have continuous and uniformly distributed marginals and that $\sigma_X(i, j) = \rho$, $\forall i \neq j$, and $\sigma_X(i, i) = 1$, where $\sigma_X(i, j)$ is the element (i, j) of Σ_X . In this case, if $\sigma_Z(i, j)$ is the element (i, j) of Σ_Z , it can be proven that [33]

$$\sigma_X(i, j) = \frac{6}{\pi} \sin^{-1} \left(\frac{\sigma_Z(i, j)}{2} \right). \quad (19)$$

In Figs. 9 and 10, we plot performance results for both the AN and NORTA cost models. We focus on a single MAC contention phase, according to Steps 2 through 4 of the SARA2 algorithm (Section 7), where the objective is to find the minimum cost node in \mathcal{S}_N . We define $c_{\min} = \min_{j \in \mathcal{S}_N} (c_j)$ and c_{winner} as the cost of the winner of the contention. Moreover, we define $\overline{\Delta c}$ and \bar{n} as the expected value of the difference $\Delta c = c_{\text{winner}} - c_{\min}$ and of the number of rounds to complete the contention, respectively. In Fig. 9, we report $\overline{\Delta c}$ as a function of the actual cost correlation ρ , considering $N = 10$ and $\Delta\rho \in \{0.05, 0.1, 0.2\}$. We observe that $N = 10$ means that we have 10 active nodes within range. This seems to be a reasonable value for the number of active neighbors in a WSN. In fact, from recent studies, such a value may be considered sufficient to have a well-connected topology. In case we had higher densities, it would make sense to put part of these nodes to sleep to save energy as too high a number of awake devices would increase the interference and reduce the network lifetime. In this figure, we depict the MAC performance accounting for the ACE, IID, and 1/N channel access functions; the resulting MAC schemes are labeled as ACE-MAC, IID-MAC, and 1/N-MAC, respectively. As expected, IID-MAC is the best scheme in terms of goodness of the solution found ($\overline{\Delta c}$), whereas 1/N-MAC is the worst. It is worth noting that ACE-MAC stays very close to IID-MAC, thereby leading to satisfactory performance. Moreover, the $\Delta\rho$ parameter can be effectively exploited to tune the ACE-MAC behavior, i.e., to decrease $\overline{\Delta c}$. Under the same settings and assumptions, in Fig. 10, we report \bar{n} as a function of the actual cost correlation ρ . In this case, the ranking between IID-MAC and 1/N-MAC is reversed. In fact, the channel

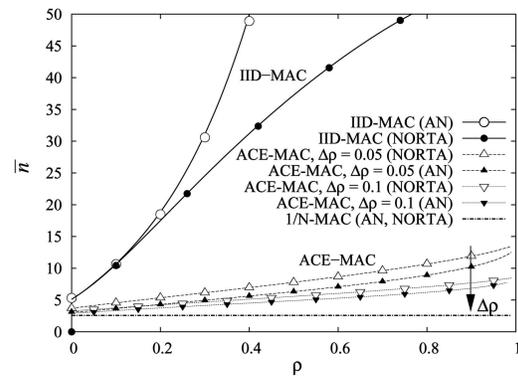


Fig. 10. Average number of iterations to complete a single contention phase. Number of nodes $N = 10$, $\Delta\rho \in \{0.05, 0.1, 0.2\}$.

access probabilities in 1/N-MAC ($1/N$) are chosen to maximize the success probability P_{succ} but completely ignoring node costs. As a consequence, 1/N-MAC performance is the best in terms of delay (\bar{n}), but performs poorly in terms of cost ($\overline{\Delta c}$). Conversely, IID-MAC penalizes high cost nodes the most (see the i.i.d. curve in Fig. 5). This explains why, within our framework, IID-MAC and 1/N-MAC are the best performing schemes for $\overline{\Delta c}$ and \bar{n} , respectively.⁹ From Fig. 10, we also stress that ACE-MAC stays sufficiently close to the lower bound (1/N-MAC) and that, once again, the $\Delta\rho$ parameter can be varied to decrease \bar{n} . As a last observation, we point out that the role of $\Delta\rho$ depends on the considered performance metric. That is, an increasing $\Delta\rho$ is beneficial for the delay ($\bar{n} \downarrow$) but has a negative impact on the goodness of the solution ($\overline{\Delta c} \uparrow$). A trade-off $\Delta\rho$ value is therefore to be found according to system/node requirements. From these graphs, it shall be observed that, even with a more general cost correlation model (NORTA), ACE-MAC still performs very close to the optimal performance. While of course not providing any proof, these results provide strong evidence that ACE-MAC is able to give very good performance without critical dependence on the cost statistics and, therefore, may be expected to have wide applicability. To sum up, neither IID-MAC nor 1/N-MAC are desirable solutions as they both provide good performance for one metric, while they perform unsatisfactorily for the other. ACE-MAC, instead, performs well for both metrics thanks to its adaptive behavior.

8.2 Performance of the Integrated MAC/Routing Schemes

In Fig. 11, we report the cumulative complementary distribution function (ccdf) of the difference between the cost (see (1)) of the path selected by the proposed *online* routing algorithms (C) and the optimal cost (C_{opt}) found by solving an offline optimization problem [35]. This cost distribution is computed by Monte Carlo simulation for the case where the hop count of the starting node is 8. In Fig. 11, we report the cost ccdf for the integrated MAC/routing SARA1 and SARA2 schemes by considering the

⁹ The exact lower bound for the delay metric in the multiple slot case is given by the access probability distribution in [21].

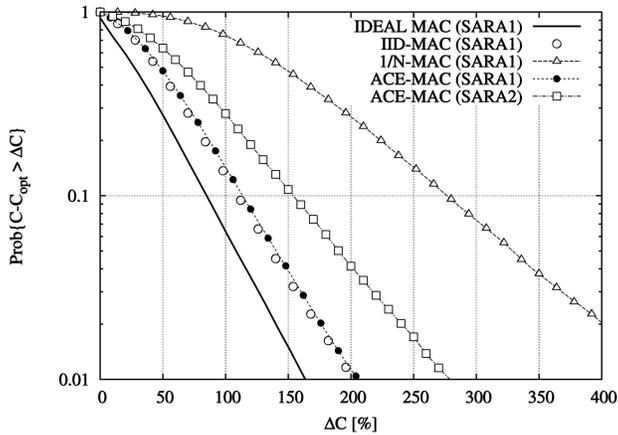


Fig. 11. Cumulative distribution function of the path cost difference between the HC integrated MAC/routing schemes and the optimal cost solution (C_{opt}). $\lambda_n = 15$, starting node hop count equal to 8, $\Delta\rho = 0.1$.

ACE-MAC, IID-MAC, and 1/N-MAC. An “IDEAL-MAC” curve, where the minimum cost nodes are always found at no communication cost, is also plotted for comparison. Node costs are uniformly distributed. Note that, when the objective is to minimize path costs, 1/N-MAC is largely suboptimal and should not be considered. On the other hand, ACE-MAC, when used with the SARA1 routing algorithm, gives satisfactory results as the degradation with respect to both IDEAL-MAC and IID-MAC is very small. It is observed that, even if IID-MAC gives the best results in terms of cost, it is not a good candidate solution as its good cost performance is achieved at the expense of a potentially very large number of contention periods (Fig. 10). Moreover, as anticipated in Section 4, we observe that SARA2 leads to worse paths in terms of cost.

In Fig. 12, we also report the performance in terms of energy expenditure. For the sake of comparison, in addition to the SARA1 and SARA2 schemes, we also report the performance of three recently proposed integrated MAC/routing approaches: GeRaF [4], NAVD [26], and IGF [22], [23]. In the graph, we report the probability that the path cost of a selected scheme will exceed the optimal cost by more than 50 percent (Ψ_{50}) as a function of the energy consumption. We assumed that every data packet consumes one unit of energy, whereas REQs, REPs, and ACKs all consume 0.1 units. The energy metric Φ is defined as the percentage of extra energy spent to advance one hop toward the sink with respect to the minimum energy expenditure (equal to 1.3, which corresponds to the sequence REQ \rightarrow REP \rightarrow DATA \rightarrow ACK). As above, the initial HC distance from the sink was considered equal to eight hops and the curves are plotted by varying the node density λ_n . Clearly, SARA1 is quite effective in promoting low cost paths. This is achieved at the expense of a slightly higher Φ , which is, however, comparable to the one obtained for the GeRaF protocol when the number of priority regions is $N_r \leq 2$. It is observed that, in our scheme, we also partition the forwarding area in two disjoint sets, $\mathcal{N}(n-1)$ and $\mathcal{N}(n)$. We can therefore assert that SARA1 does not introduce a larger one-hop energy cost with respect to GeRaF when we subdivide the forwarding area into the same number of regions (2). This is due to the high efficiency of our MAC, where high cost nodes are ruled out

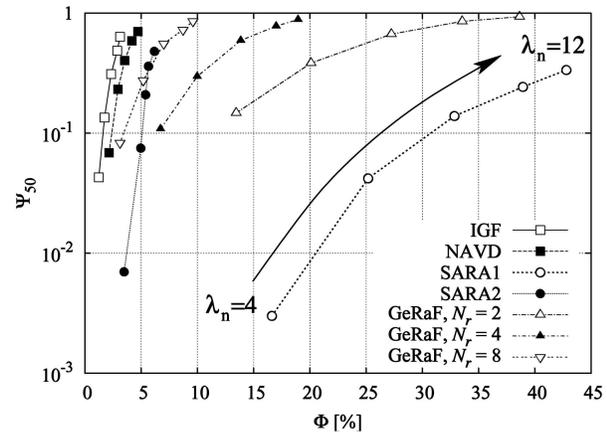


Fig. 12. Trade-off between path quality (Ψ_{50}) and one-hop energy cost (Φ).

right in the channel access phase. As expected, for an increased N_r GeRaF is more efficient as the number of nodes that will contend for the channel within the same round will be further reduced. As far as SARA2 is concerned, we observe that its energy consumption is very low ($\Phi \leq 5\%$) and comparable to the other schemes. In practice, when the priority is to select cost efficient paths (e.g., to balance traffic and node energies, promote reliable paths, etc.), SARA1 may be a good alternative (Fig. 11). If, instead, we want to minimize the energy consumption, then we may want to consider the SARA2 algorithm as its energy consumption is still comparable to the solutions proposed so far in the literature, whereas its path cost is significantly lower. Hence, whether to use SARA1 or SARA2 actually depends on the trade-off between energy and cost.

9 CONCLUSIONS

In this work, we addressed routing and MAC algorithms for the efficient delivery of packets in wireless sensor networks. We first discussed *online* algorithms to efficiently route packets over hop count coordinates. The goal of these routing schemes is to minimize a generic cost function that depends on the node costs encountered along the path and that can be related to various metrics, such as node residual energies or queue lengths. Subsequently, we proposed a MAC technique to be coupled with the above routing algorithms. This MAC is specifically designed to efficiently locate, at a low communication cost, the minimum cost nodes within coverage. This information is then exploited by *online* routing algorithms to make forwarding decisions. In the last part of the paper, we reported performance results as evidence of the effectiveness of the proposed approaches.

The material treated in the present paper is a step toward the definition of efficient, integrated MAC/routing schemes for wireless sensor networks. Differently from the approaches proposed so far, we designed our schemes to work over HC topologies rather than considering geographical routing. These algorithms may, in fact, be useful in scenarios where the location information is not available. We stress that a comparison with stateful routing schemes is not addressed here as this would require making specific assumptions about the exact network dynamics, which goes

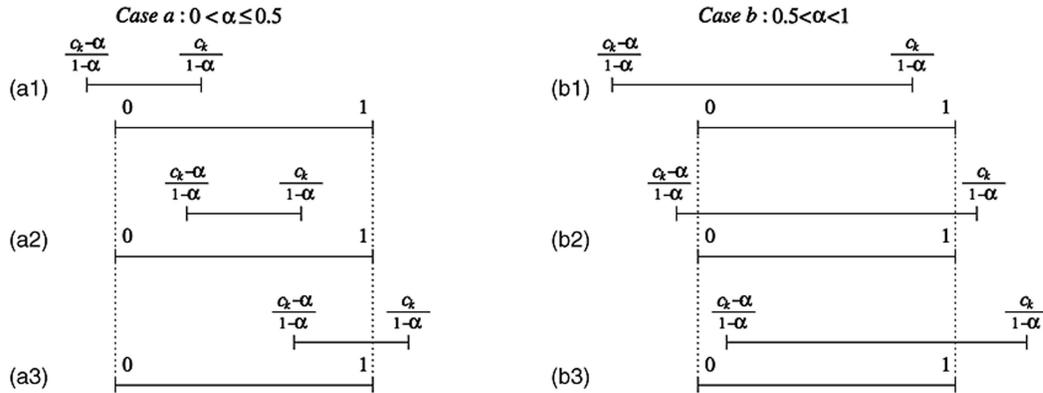


Fig. 13. Diagram for the calculation of the convolution integral in (22). The intervals represent the values of \bar{c} for which $f(c_k|\bar{c})$ ($\bar{c} \in [(c_k - \alpha)/(1 - \alpha), c_k/(1 - \alpha)]$) and $f_{\bar{c}}(\bar{c})$ ($\bar{c} \in [0, 1]$) are different from zero.

beyond the scope of the present paper. However, we note that, for any contention-based (stateful) MAC protocol comparable to the one considered here, we must include the additional signaling burden related to topology information exchanges among neighbors for routing table maintenance, which may lead to worse performance in the highly dynamic environments we considered in this paper. A more precise characterization of this trade-off, as well as the consideration of possible hybrid schemes, is currently under study. Also, the material that we presented in this paper motivates further research. Our current work includes issues related to the estimation of the local density, further optimizations of the joint MAC/routing solution and the possibility of coupling our approach with data fusion techniques. In fact, given its flexibility, the proposed MAC can be easily adapted to drive the contention/next hop selection procedures toward the most profitable paths by accounting for data fusion rules.

APPENDIX

Consider the cost model introduced in Section 6.2 and consider a generic set of N nodes \mathcal{S}_N . In this Appendix, we will derive the probability that the generic node $k \in \mathcal{S}_N$ is the minimum cost node in the set ((12), (13), and (14)). The actual cost at node k is $c_k = \bar{c} + \gamma_k$, where the common cost component \bar{c} is uniformly distributed in $[0, 1]$ and γ_k is uniformly distributed in $[-\alpha\bar{c}, \alpha(1 - \bar{c})]$. If Γ_k is the r.v. associated with γ_k , we have that $f_{\Gamma}(\gamma|\bar{c})$, the conditional probability density function (pdf) of Γ_k , is

$$f_{\Gamma}(\gamma|\bar{c}) = \begin{cases} \alpha^{-1} & \gamma \in [-\alpha\bar{c}, \alpha(1 - \bar{c})] \\ 0 & \text{elsewhere.} \end{cases} \quad (20)$$

Moreover, considering the cost c_k as given, the pdf of \bar{c} given c_k , $f(\bar{c}|c_k)$ can be obtained via the Bayes rule as follows: $f(\bar{c}|c_k) = f(c_k|\bar{c})f_{\bar{c}}(\bar{c})/f_C(c_k)$, where $f(c_k|\bar{c})$ is derived from (20) through a domain shift

$$f(c_k|\bar{c}) = \begin{cases} \alpha^{-1} & c_k \in [\bar{c} - \alpha\bar{c}, \bar{c} + \alpha(1 - \bar{c})] \\ 0 & \text{elsewhere.} \end{cases} \quad (21)$$

$f_C(c_k)$, the pdf of the cost at node c_k , is found via the following convolution integral:

$$f_C(c_k) = \int_{-\infty}^{+\infty} f_{\bar{c}}(\bar{c}) f_{\Gamma}(c_k - \bar{c}|\bar{c}) d\bar{c}, \quad (22)$$

where the pdf associated with the common cost component $f_{\bar{c}}(\bar{c})$ is uniform in $[0, 1]$ and $f_{\Gamma}(\gamma_k|\bar{c})$ is defined as in (20). From the definition intervals of $f_{\bar{c}}(\cdot)$ and $f_{\Gamma}(\cdot)$, expressed as a function of \bar{c} (the common cost part), one can calculate the integral in (22) for six disjoint cases. These cases are reported in Fig. 13. In fact, the length of the definition interval of $f_{\Gamma}(\gamma_k|\bar{c})$ as a function of \bar{c} (the interval is $[(c_k - \alpha)/(1 - \alpha), c_k/(1 - \alpha)]$) is given by $\alpha/(1 - \alpha)$ and is smaller than or equal to 1 for $0 < \alpha \leq 0.5$ and larger than 1 when $0.5 < \alpha < 1$; this gives the two cases *a* and *b* in Fig. 13. Moreover, for each of these cases, we have three possibilities, depending on the value of c_k . Hence, by solving the integral above, one achieves two different expressions for $f_C(c_k)$, according to the value of α . In particular, $f_C(c_k)$ is found for the two cases, $0 < \alpha \leq 0.5$ (Case *a*) and $0.5 < \alpha < 1$ (Case *b*), as follows:

$$f_C(c_k) \Big|_{0 < \alpha \leq 0.5} = \begin{cases} \frac{c_k}{\alpha(1-\alpha)} & 0 \leq c_k < \alpha \\ \frac{1}{1-\alpha} & \alpha \leq c_k \leq 1 - \alpha \\ \frac{1-c_k}{\alpha(1-\alpha)} & 1 - \alpha < c_k \leq 1 \\ 0 & \text{elsewhere,} \end{cases} \quad (23)$$

$$f_C(c_k) \Big|_{0.5 < \alpha < 1} = \begin{cases} \frac{c_k}{\alpha(1-\alpha)} & 0 \leq c_k < 1 - \alpha \\ \alpha^{-1} & 1 - \alpha \leq c_k \leq \alpha \\ \frac{1-c_k}{\alpha(1-\alpha)} & \alpha < c_k \leq 1 \\ 0 & \text{elsewhere.} \end{cases} \quad (24)$$

The conditional pdf $f(\bar{c}|c_k)$ can therefore be obtained as $f(\bar{c}|c_k) = f(c_k|\bar{c})f_{\bar{c}}(\bar{c})/f_C(c_k)$, which can be written in compact form as

$$f(\bar{c}|c_k) = \begin{cases} (\alpha f_C(c_k))^{-1} & \bar{c} \in \mathcal{I}_{\alpha} \\ 0 & \text{elsewhere,} \end{cases} \quad (25)$$

where $f_C(c_k)$ is one of the two expressions above, according to the value of α . The interval $\mathcal{I}_{\alpha} = [\bar{c}_{min}, \bar{c}_{max}]$ is specified by

$$\bar{c}_{min} = \max\left(0, \frac{c_k - \alpha}{1 - \alpha}\right), \quad \bar{c}_{max} = \min\left(1, \frac{c_k}{1 - \alpha}\right). \quad (26)$$

Now, we refer to $P_{min}^k(c_k)$ as the probability that node k with cost c_k is the minimum cost node in \mathcal{S}_N . $P_{min}^k(c_k)$ is given by

$$P_{min}^k(c_k) = \int_{\bar{c} \in \mathcal{I}_a} f(\bar{c}|c_k) (1 - F_{\Gamma}(c_k - \bar{c}))^{N-1} d\bar{c}, \quad (27)$$

where $F_{\Gamma}(\cdot)$ is the conditional cdf of the r.v. Γ_k for the cost displacement at node k . The above equation returns the probability that all of the remaining $N - 1$ nodes in \mathcal{S}_N have a cost which is higher than or equal to c_k (term $(1 - F_{\Gamma}(c_k - \bar{c}))^{N-1}$), given that the cost of node k is c_k . The above integral must be evaluated considering the six cases above. For the sake of illustration, in the following, we report the calculation for Case a :

- a1. For a given value of $\alpha \in (0, 0.5]$, the integral in (27) is significant when $f(\bar{c}|c_k) > 0$. Moreover, $f(\bar{c}|c_k) > 0$ when both $f(c_k|\bar{c})$ and $f_{\bar{c}}(\bar{c})$ are larger than zero and this happens when $0 \leq \bar{c} < c_k/(1 - \alpha)$ (see Case $a1$ in Fig. 13). This gives the integration limits for (27). The pdf $f_C(c_k)$ that has to be considered to derive $f(\bar{c}|c_k)$ is given by the first line in (23). In fact, Case $a1$ is equivalent to having $(c_k - \alpha)/(1 - \alpha)$, the lower limit of the interval on the left-top side of Fig. 13, smaller than zero, which translates to $0 \leq c_k < \alpha$, therefore leading to the expression for $f_C(c_k) = c_k/(\alpha(1 - \alpha))$. Hence, $f(\bar{c}|c_k) = (1 - \alpha)/c_k$ (see (25)), the cdf $F_{\Gamma}(\gamma) = (\gamma + \alpha\bar{c})/\alpha$, and (27) results in

$$\begin{aligned} & \int_0^{c_k/(1-\alpha)} \frac{(\alpha - c_k + \bar{c} - \alpha\bar{c})^{N-1} (1 - \alpha)}{c_k \alpha^{N-1}} d\bar{c} = \\ & = \frac{1}{c_k \alpha^{N-1}} \int_{\alpha - c_k}^{\alpha} \xi^{N-1} d\xi = \frac{\alpha^N - (\alpha - c_k)^N}{c_k N \alpha^{N-1}} \end{aligned} \quad (28)$$

and this gives the first line in (12), which holds for $0 \leq c_k < \alpha$ due to the condition on $f_C(c_k)$.

- a2. For a given value of $\alpha \in (0, 0.5]$ and the reasoning above, the interval to be considered for \bar{c} is $(c_k - \alpha)/(1 - \alpha) \leq \bar{c} \leq c_k/(1 - \alpha)$ and the region for c_k ($f_C(c_k)$) is $\alpha \leq c_k \leq 1 - \alpha$, which comes from the fact that, in Case $a2$, $(c_k - \alpha)/(1 - \alpha) \geq 0$ and $c_k/(1 - \alpha) \leq 1$, respectively. Hence, $f_C(c_k) = 1/(1 - \alpha)$ and $f(\bar{c}|c_k) = (1 - \alpha)/\alpha$. Equation (27) results in

$$\int_{(c_k - \alpha)/(1 - \alpha)}^{c_k/(1 - \alpha)} \frac{(\alpha - c_k + \bar{c} - \alpha\bar{c})^{N-1} (1 - \alpha)}{\alpha^N} d\bar{c} = \frac{1}{N} \quad (29)$$

and this gives the second line in (12), which holds for $\alpha \leq c_k \leq 1 - \alpha$ due to the condition on $f_C(c_k)$.

- a3. For a given value of $\alpha \in (0, 0.5]$, \bar{c} now spans in the range $(c_k - \alpha)/(1 - \alpha) \leq \bar{c} \leq 1$ and the condition for c_k ($f_C(c_k)$) is $1 - \alpha < c_k \leq 1$. Hence, $f_C(c_k) = (1 - c_k)/(\alpha(1 - \alpha))$ and $f(\bar{c}|c_k) = (1 - \alpha)/(1 - c_k)$. Equation (27) results in

$$\begin{aligned} & \int_{(c_k - \alpha)/(1 - \alpha)}^1 \frac{(\alpha - c_k + \bar{c} - \alpha\bar{c})^{N-1} (1 - \alpha)}{(1 - c_k) \alpha^{N-1}} d\bar{c} \\ & = \frac{(1 - c_k)^{N-1}}{N \alpha^{N-1}} \end{aligned} \quad (30)$$

and this gives the last line in (12), which holds for $1 - \alpha < c_k \leq 1$, due to the condition on $f_C(c_k)$.

The derivation for Case b is obtained similarly, considering the diagram for Cases $b1$, $b2$, and $b3$ in Fig. 13.

ACKNOWLEDGMENTS

Part of this work was presented at the Third International Workshop on Measurement, Modeling, and Performance Analysis of Wireless Sensor Networks (SenMetrics '05). This material is based upon work partially supported by the European Commission under contract IST-4-027227-IP (eSENSE).

REFERENCES

- [1] I.F. Akyildiz, W. Su, Y. Sankarasubramanian, and E. Cayirci, "A Survey on Sensor Networks," *IEEE Comm. Magazine*, vol. 40, no. 8, pp. 102-116, Aug. 2002.
- [2] F. Ye, A. Chen, S. Lu, and L. Zhang, "A Scalable Solution to Minimum Cost Forwarding in Large Sensor Networks," *Proc. IEEE 10th Int'l Conf. Computer Comm. and Networks*, pp. 304-309, 2001.
- [3] F. Ye, G. Zhong, S. Lu, and L. Zhang, "GRADient Broadcast: A Robust Data Delivery Protocol for Large Scale Sensor Networks," *ACM Wireless Networks*, vol. 11, no. 2, pp. 285-298, May 2005.
- [4] M. Zorzi and R.R. Rao, "Geographic Random Forwarding (GeRaF) for Ad Hoc and Sensor Networks: Multihop Performance," *IEEE Trans. Mobile Computing*, vol. 2, no. 4, pp. 337-348, Oct.-Dec. 2003.
- [5] T. Melodia, D. Pompili, and F.I. Akyildiz, "Optimal Location Topology Knowledge for Energy Efficient Geographical Routing in Sensor Networks," *Proc. IEEE INFOCOM*, Mar. 2004.
- [6] Z. Haas, "Design Methodologies for Adaptive and Multimedia Networks," *IEEE Comm. Magazine*, vol. 39, no. 11, pp. 106-107, Nov. 2001.
- [7] G. Dimiç, N.D. Sidiropoulos, and R. Zhang, "Medium Access Control-Physical Cross-Layer Design," *IEEE Signal Processing Magazine*, pp. 40-50, Sept. 2004.
- [8] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, "Directed Diffusion for Wireless Sensor Networking," *IEEE/ACM Trans. Networking*, vol. 1, pp. 2-16, Feb. 2003.
- [9] I. Stojmenovic and X. Lin, "Loop-Free Hybrid Single-Path/Flooding Routing Algorithms with Guaranteed Delivery for Wireless Networks," *IEEE Trans. Parallel and Distributed Systems*, vol. 12, no. 10, Oct. 2001.
- [10] B. Karp and H.T. Kung, "GPSR: Greedy Perimeter Stateless Routing for Wireless Networks," *Proc. IEEE/ACM MobiCom*, pp. 243-254, Aug. 2000.
- [11] Y. Xu and J. Heidemann, "Geographic-Informed Energy Conservation for Ad Hoc Routing," *Proc. IEEE/ACM MobiCom*, pp. 70-77, July 2001.
- [12] P. Bose, P. Morin, I. Stojmenović, and J. Urrutia, "Routing with Guaranteed Delivery in Ad Hoc Wireless Networks," *ACM Wireless Networks*, vol. 7, pp. 609-616, 2001.
- [13] G. Xing, C. Lu, R. Pless, and Q. Huang, "On Greedy Geographic Routing Algorithms in Sensing-Covered Networks," *Proc. Fifth ACM Int'l Symp. Mobile Ad Hoc Networking and Computing*, pp. 31-42, 2004.
- [14] F. Kuhn, R. Wattenhofer, Y. Zhang, and A. Zollinger, "Geometric Ad-Hoc Routing: Of Theory and Practice," *Proc. Ann. ACM Symp. Principles of Distributed Computing*, pp. 63-72, 2003.
- [15] L. Kleinrock and J. Silvester, "Optimum Transmission Radii for Packet Radio Networks or Why Six Is a Magic Number," *Proc. IEEE Nat'l Telecomm. Conf.*, pp. 4.3.1-4.3.5, 1978.
- [16] H. Takagi and L. Kleinrock, "Optimal Transmission Ranges for Randomly Distributed Radio Packet Terminals," *IEEE Trans. Comm.*, vol. 32, no. 3, pp. 246-257, 1984.
- [17] G. Finn, "Routing and Addressing Problems in Large Metropolitan Scale Internetworks," Technical Report ISI/RR-87-180, Information Science Inst., Mar. 1987.
- [18] C.E. Perkins, E.M. Belding-Royer, and S.R. Das, "Ad-Hoc On Demand Distance Vector Routing," IETF Mobile Ad-Hoc Networks (MANET) Working Group RFC, <http://www.ietf.org/rfc/rfc3561.txt>, July 2003.

- [19] D.B. Johnson and D.A. Maltz, "Dynamic Source Routing in Ad Hoc Wireless Networks," *Mobile Computing*, T. Imielinsky and H. Korth, eds., chapter 5, pp. 153-181, Kluwer, 1996.
- [20] B.N. Karp, "Geographic Routing for Wireless Networks," PhD thesis, Harvard Univ., 2000.
- [21] Y.C. Tay, K. Jamieson, and H. Balakrishnan, "Collision-Minimizing CSMA and Its Applications to Wireless Sensor Networks," *IEEE J. Selected Areas in Comm.*, vol. 22, no. 6, pp. 1048-1057, Aug. 2004.
- [22] B. Blum, T. He, S. Son, and J. Stankovic, "IGF: A Robust State-Free Communication Protocol for Sensor Networks," Technical Report CS-2003-11, Computer Science Dept., Univ. of Virginia, 2003.
- [23] D. Chen, J. Deng, and P.K. Varshney, "A State-Free Data Delivery Protocol for Multihop Wireless Sensor Networks," *Proc. Wireless Comm. and Networking Conf.*, 2005.
- [24] M. Rossi and M. Zorzi, "Cost Efficient Localized Geographical Forwarding Strategies for Wireless Sensor Networks," *Distributed Cooperative Laboratories: Networking, Instrumentation and Measurements*, F. Davoli, S. Palazzo, and S. Zappatore, eds., 2006.
- [25] X. Jinghao, B. Peric, and B. Vojcic, "Energy-Aware and Link-Adaptive Routing Metrics for Ultra Wideband Sensor Networks," *Proc. IEEE Workshop Ultra Wide Band for Sensor Networks*, July 2005.
- [26] S. Lee, B. Bhattacharjee, and S. Banerjee, "Efficient Geographic Routing in Multihop Wireless Networks," *Proc. ACM MobiHoc '05*, May 2005.
- [27] R. Guérin and A. Orda, "Computing Shortest Paths for Any Number of Hops," *IEEE/ACM Trans. Networking*, vol. 10, no. 5, pp. 613-620, Oct. 2002.
- [28] M. Rossi, M. Zorzi, and R.R. Rao, "Statistically Assisted Routing Algorithms (SARA) for Hop Count Based Forwarding in Wireless Sensor Networks," *ACM/Springer J. Wireless Networks*, to appear.
- [29] S.M. Ross, *Introduction to Stochastic Dynamic Programming*. Academic Press, 1983.
- [30] P. Popovski, F.H. Fitzek, and R. Prasad, "Batch Conflict Resolution Algorithm with Progressively Accurate Multiplicity Estimation," *Proc. ACM Workshop Discrete Algorithms and Methods for Mobile Computing and Comm. (DIAL M-POMC '04)*, pp. 31-40, Oct. 2004.
- [31] G.D. Bacco, T. Melodia, and F. Cuomo, "A MAC Protocol for Delay-Bounded Applications in Wireless Sensor Networks," *Proc. Third Ann. Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net '04)*, June 2004.
- [32] B.N. Clark, C.J. Colbourn, and D.S. Johnson, "Unit Disk Graphs," *Discrete Math.*, vol. 86, nos. 1-3, pp. 165-177, Aug. 1991.
- [33] M. Cario and B. Nelson, "Modeling and Generating Random Vectors with Arbitrary Marginal Distribution and Correlation Matrix," technical report, Dept. of Industrial Eng. and Management Sciences, Northwestern Univ., 1997.
- [34] S. Ghosh and S.G. Henderson, "Properties of the NORTA Method in Higher Dimensions," *Proc. 2002 Winter Simulation Conf.*, Dec. 2002.
- [35] W.-K. Chen, *Theory of Nets: Flow in Networks*. John Wiley & Sons, 1990.



Michele Rossi received the Laurea degree in electrical engineering (with honors) and the PhD degree in information engineering from the University of Ferrara in 2000 and 2004, respectively. Since 2000, he has been a research fellow in the Department of Engineering at the University of Ferrara. During 2003, he was on leave at the Center for Wireless Communications (CWC) at the University of California, San Diego (UCSD), where he did research on wireless sensor networks. In November 2005, he joined the Department of Information Engineering at the University of Padova, Italy, where he is currently an assistant professor. Dr. Rossi is currently part of the EU-funded Ambient Networks and eSENSE projects. His research interests include TCP/IP protocols over wireless networks, performance analysis of link layer retransmission techniques, routing and access selection in heterogeneous wireless networks, integrated MAC/routing algorithms, and in-network data aggregation techniques for wireless sensor networks. He is a member of the IEEE.



Michele Zorzi received the Laurea degree and the PhD degree in electrical engineering from the University of Padova, Italy, in 1990 and 1994, respectively. During the academic year of 1992/93, he was on leave at the University of California, San Diego (UCSD), attending graduate courses and doing research on multiple access in mobile radio networks. In 1993, he joined the faculty of the Dipartimento di Elettrotecnica e Informazione, Politecnico di Milano, Italy. After spending three years with the Center for Wireless Communications at UCSD, in 1998, he joined the School of Engineering at the University of Ferrara, Italy, and, in 2003, joined the Department of Information Engineering of the University of Padova, Italy, where he is currently a professor. His present research interests include performance evaluation in mobile communications systems, random access in mobile radio networks, ad hoc and sensor networks, and energy constrained communications protocols. From 2003 to 2005, Dr. Zorzi was the editor-in-chief of the *IEEE Wireless Communications Magazine* and currently serves on the editorial boards of the *IEEE Transactions on Communications*, the *IEEE Transactions on Wireless Communications*, the *IEEE Transactions on Mobile Computing*, the *Wiley Journal of Wireless Communications and Mobile Computing*, and the *ACM/URSI/Kluwer Journal of Wireless Networks*. He was also a guest editor for special issues of the *IEEE Personal Communications Magazine* (energy management in personal communications systems) and the *IEEE Journal on Selected Areas in Communications* (multimedia network radios). He is a fellow of the IEEE.

► For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.