Abstract—In this paper we address the issue of the optimal candidate-set selection in the opportunistic routing paradigm. More specifically, although several algorithms for selecting the optimal candidate set have been proposed, to the best of our knowledge none of them has never considered the problem of selecting the optimal constrained candidate set, namely the optimal candidate set with a fixed maximum set size. In this paper we contribute to this problem by providing an analytical framework to model both the optimal constrained and unconstrained candidate-set selection. Moreover, we propose two algorithms for optimal candidate-set selection for distance vector routing, one for the constrained and one for the unconstrained case. Simulations based on experimental data validate our proposal.

Index Terms—ad hoc networks, opportunistic routing, candidate selection, constrained candidate set size.

I. INTRODUCTION

In the last years, the opportunistic routing paradigm [1]–[3] has been proposed to exploit the broadcast nature of the wireless communications to compensate the channel unreliability. According to this paradigm, a forwarder simply broadcasts the data packet to a set of nodes, namely a candidate set. The next-hop selection is performed at the receiver side by choosing, among the candidates that correctly received the packet, the best one according to a certain routing progress metric.

Very recently, the problem of selecting the candidate set which assures the highest routing progress, namely the optimal candidate-set selection problem, has received huge attention. However, the existing algorithms for optimal candidate-set selection assume an unconstrained set size. Since the set size deeply affects the routing performances in terms of overhead for candidate coordination, the problem of selecting the optimal candidate set with a fixed maximum set size (optimal constrained candidate-set selection) has a notable relevance and it is still an open problem.

In this paper we contribute to such a problem by proposing an analytical framework to model both the constrained and unconstrained candidate-set selection. Moreover, we propose two algorithms for optimal candidate-set selection for distance vector routing, one for the constrained and one for the unconstrained case. As regard to the constrained selection, the proposed algorithm exhibits an exponential time-complexity rather than the factorial complexity associated with the exhaustive search. It is worthwhile to underline that the proposed algorithm for unconstrained selection has been considered for the sake of comparison with respect to the constraint one.

The paper is organized as follows: in Section II we present the related works, whereas in Section III we introduce the adopted analytical model. Section IV presents the proposed algorithms along with the proof of their optimality. In Section V a performance evaluation based on experimental data is illustrated and, finally, in the last section some conclusions are drawn.

II. RELATED WORKS

Several works address the problem of optimal unconstrained candidate-set selection but none of them deals with the constrained problem.

In [4] the authors are the first to propose a routing metric, the expected anypath number of transmissions (EAX), which accounts for the specific characteristics of opportunistic paradigm. They present also a distance vector algorithm for the unconstrained selection problem with a polynomial complexity equal to $O(n^2)$ where $n$ is the neighbor set size, without proving its optimality. Several other works propose different heuristics for unconstrained candidate-set selection [5], [6].

In [7] the authors generalize the Bellman-Ford algorithm for the unconstrained problem with a complexity of $O(2^n)$ and they prove its optimality. Moreover, the authors provide an algorithm with linearithmic complexity $O(n \log n)$ and a proof of its optimality for some special cases.

In [8] the authors propose a link state algorithm for the unconstrained selection based on the Dijkstra algorithm. In [9] the Dijkstra algorithm is generalized to the opportunistic routing paradigm exhibiting the same linearithmic complexity of the shortest-path algorithm.
III. ANALYTICAL MODEL

We model a wireless network with a probabilistic direct graph [10]:

\[ G = (V, L, D) \]  

(1)

in which a vertex \( v \in V \) denotes a wireless node and an edge \( l_{u,v} \in L \) represents a communication link from node \( u \) to node \( v \). Each link \( l_{u,v} \) is characterized by a delivery ratio \( d_{u,v} \in D \), which measures the probability that a packet is correctly received in a single transmission along such a link.

We suppose that a node \( s \in V \) has to forward a packet toward a destination \( d \in V \) and, without loss of generality, we assume that the link failure events are statistically independent of each other [4], [8], [9].

We define as \( N \subseteq V \) the set of neighbors of \( s \). Since the neighbors of \( s \) are the candidates to become the next relay for the packet, we define as candidate set \( C \) for the flow \( (s, d) \) an ordered sequence of neighbors:

\[ C = \{(v_1, v_2, \ldots, v_m) : v_i \neq v_j, v_i \in N\}. \]  

(2)

The order among the candidates reflects the priorities for packet retransmission, i.e., candidate \( v_i \) is allowed to become the next forwarder if only if all the candidates with higher priority \( v_1, \ldots, v_{i-1} \) have failed. The probability \( p_{v_i} \) that the candidate \( v_i \) becomes the next relay for the packet is\(^1\):

\[ p_{v_i} = d_i \prod_{j=1}^{i-1} (1 - d_j) = d_i \prod_{j=1}^{i-1} 1 - d_j \]  

(3)

and the probability \( p_s \) that \( s \) has to retransmit the packet is

\[ p_s = 1 - \sum_{v_i \in C} p_{v_i} = \prod_{i=1}^{m} d_i \]  

(4)

Assuming that the cardinality of the neighbor set is \( |N| = n \), the number of available candidate sets is equal to\(^2\):

\[ \sum_{k=1}^{n} \frac{n!}{(n-k)!} = e \Gamma(n + 1, 1) - 1 \approx e n! \]  

(5)

where \( \frac{n!}{(n-k)!} \) is the number of ordered \( k \)-permutations of \( n \) elements without repetition and \( \Gamma(\cdot, \cdot) \) is the upper incomplete gamma function.

The optimum candidate set \( C_{opt} \) is the set that maximizes the progress toward the destination by minimizing the distance of \( s \) from \( d \) measured according to the routing metric.

In the following, we adopt as routing metric \( f(\cdot) \) the average number of opportunistic data-link transmissions (EOTX) [4], [10], but we underline that the proposed analytical model can be easily extended to other routing metrics.

\(^1\)In the following, we define \( d_i = d_{s,v_i} \) to simplify the notation.

\(^2\)After some algebraic manipulation and using the notable relation

\[ \Gamma(s, x) = (s - 1)! e^{-x} \sum_{k=0}^{s-1} \frac{x^k}{k!} \]

If we indicate with \( C \) the set of the admissible candidate sets, the optimum candidate set \( C_{opt} \) is defined as:

\[ f(C_{opt}) = \min_{C \subseteq C} f(C) \]  

(6)

where \( f(C) \), namely the EOTX needed to deliver a packet toward the candidate set \( C \), is equal to [4], [10]:

\[ f(C) = \frac{1}{1 - p_s} \left( 1 + \sum_{i=1}^{m} p_{v_i} f_{v_i} \right) \]  

(7)

In (7) \( f_{v_i} \) is the minimum EOTX for the flow \((v_i, d)\), i.e. the EOTX toward the optimum candidate set for the flow \((v_i, d)\).

IV. OPTIMAL CANDIDATE SET SELECTION

In this section we present the general theory for optimal candidate-set selection, and then we describe the proposed algorithms for both optimal constrained and unconstrained candidate-set selection.

A. Theory of Optimal Candidate Set Selection

In the following, we first introduce a theorem that states the priority rule among the candidates and that allows us to notably reduce the complexity of the candidate selection for both the constrained and unconstrained problems. Then, we present a proof different from the one given in [9] for optimal unconstrained candidate selection rule that reduces the complexity to a linearithmic factor. Finally, we prove that the mentioned priority rule does not hold in the case of constrained selection.

**Theorem 1 (Priority Rule):** Let be \((s, d)\) a generic communication flow. Any candidate set, ordered according to the EOTX of each candidate, achieves a EOTX (needed to deliver a packet toward the destination) lower than the one associated with a candidate set composed by the same candidates but in a different order.

**Proof:** Let be \( C = \{v_1, \ldots, v_k, v_l, v_m, \ldots, v_s, v_l\} \) the candidate set ordered according to the minimum EOTX, namely, \( f_{v_1} < \ldots < f_{v_k} < f_{v_l} < \ldots < f_{v_s} \). Let be \( C' = \{v_1, \ldots, v_k, v_m, v_l, \ldots, v_s, v_l\} \) a different candidate set formed by the same candidates but ordered according to a different criterion. We want to show that

\[ f(C) < f(C') \]  

(8)

In order to prove the theorem we conduct a reduction ad absurdum, i.e. we suppose that

\[ f(C) > f(C') \]  

(9)

Basing on (7), we can write \( f(C) \) and \( f(C') \) respectively as in (10) and (11) [10] reported in the next page. Combining (10) and (11) in (9) and simplifying the equal terms, one has:

\[ f_{v_l} d_l \prod_{z=1}^{k} \bar{d}_z = f_{v_m} d_m \prod_{z=1}^{k} \bar{d}_z > f_{v_u} d_u \prod_{z=1}^{k} \bar{d}_z + f_{v_l} d_l \bar{d}_l \prod_{z=1}^{k} \bar{d}_z \]

\[ \iff f_{v_l} d_l + f_{v_m} d_m \bar{d}_l > f_{v_u} d_u + f_{v_l} d_l \bar{d}_l \]  

(12)
\[
f(C) = f(v_1, \ldots, v_k, v_m, \ldots, v_s, v_t) = \frac{1 + f_{v_1} d_1 + \ldots + f_{v_k} d_k \prod_{i=1}^{k} \tilde{d}_1 + f_{v_m} d_m \prod_{i=1}^{m} \tilde{d}_2 + \ldots + f_{v_s} d_s \prod_{i=1}^{s} \tilde{d}_z}{1 - \prod_{i=1}^{k} \tilde{d}_z} \tag{10}
\]

\[
f(C') = f(v_1, \ldots, v_k, v_m, \ldots, v_s, v_t) = \frac{1 + f_{v_1} d_1 + \ldots + f_{v_m} d_m \prod_{i=1}^{m} \tilde{d}_2 + f_{v_1} d_1 \prod_{i=1}^{1} \tilde{d}_z + \ldots + f_{v_s} d_s \prod_{i=1}^{s} \tilde{d}_z}{1 - \prod_{i=1}^{m} \tilde{d}_z} \tag{11}
\]

from which we conclude that:
\[
f(C) > f(C') \iff f_{v_1} > f_{v_m} \tag{13}
\]

The equation (13) is an absurd since for hypothesis \(f_{v_1} < f_{v_m}\).

**Proposition 1:** Let us consider a neighbor set \(\mathbb{N}\) with cardinality \(|\mathbb{N}| = n\) ordered according to the EOTX metric, namely \(f_{v_1} < f_{v_{i+1}}, \forall i < n\). The node \(v_1\) belongs to the optimal unconstrained candidate set if and only if the cardinality of the optimal set is greater or equal than \(i\):
\[
v_1 \in C_{opt} \iff |C_{opt}| \geq i \tag{14}
\]

**Proof:** We conduct a reductio ad absurdum, i.e. we suppose that \(v_1 \notin C_{opt}\). As consequence, one has:
\[
\min_{C \in \mathcal{E}} f(C) \leq f(C_{opt}) = f(v_1, C_{opt}) \tag{15}
\]

Accounting for (7) and after some algebraic manipulations, \(f(v_1, C_{opt})\) can be written as:
\[
f(v_1, C_{opt}) = \frac{(1 + f_{v_1} d_1 + \tilde{d}_1 f(C_{opt})(1 - \prod_{v \in C_{opt}} \tilde{d}_1)}{1 - \tilde{d}_1 \prod_{v \in C_{opt}} \tilde{d}_1} \tag{16}
\]

Combining (16) in (15) and taking into account that the denominator of (16) is a positive quantity, we obtain that:
\[
f(C_{opt})(1 - \tilde{d}_1 \prod_{v \in C_{opt}} \tilde{d}_1) < (1 + f_{v_1} + \tilde{d}_1 f(C_{opt})(1 - \prod_{v \in C_{opt}} \tilde{d}_1) \tag{17}
\]

Since
\[
1 - \tilde{d}_1 \prod_{v \in C_{opt}} \tilde{d}_1 = d_1 + \tilde{d}_1(1 - \prod_{v \in C_{opt}} \tilde{d}_1) \tag{18}
\]

by substituting (18) in (17), after simple algebraic manipulations, one has:
\[
f(C_{opt}) < 1 + f_{v_1} \tag{19}
\]

If we indicate with \(C_{opt} = \{v_j, v_k, \ldots, v_s, v_m\}\), (19) is equivalent to:
\[
\frac{1 + f_{v_j} d_j + f_{v_k} d_k \tilde{d}_j + \ldots + f_{v_m} d_m \tilde{d}_m \prod_{z=k}^{m} \tilde{d}_z}{1 - \tilde{d}_j \prod_{z=k}^{m} \tilde{d}_z} < 1 + f_{v_1} \tag{20}
\]

Taking into account that \(1 - \tilde{d}_j \prod_{z=k}^{m} \tilde{d}_z = d_j + \tilde{d}_j d_k + \ldots + \tilde{d}_j d_m \prod_{z=k}^{m} \tilde{d}_z\), the equation (20) can be rewritten as:
\[
(f_{v_j} - f_{v_1}) d_j + (f_{v_k} - f_{v_1}) d_k \tilde{d}_j + \ldots + \tilde{d}_j m \prod_{z=k}^{m} \tilde{d}_z < 0 \tag{21}
\]

that constitutes an absurd since the left term of the inequality is positive due to the assumption \(f_{v_j} < f_{v_1}, \forall v_i \in C_{opt}\), whereas the right one is negative.

**Proposition 2:** Let us consider a neighbor set \(\mathbb{N}\) with cardinality \(|\mathbb{N}| = n\) ordered according to the EOTX metric, namely \(f_{v_1} < f_{v_{i+1}}, \forall i < n\). The node \(v_1\) belongs to the optimal unconstrained candidate set if and only if the cardinality of the optimal set is greater or equal than \(i\):
\[
v_1 \in C_{opt} \iff |C_{opt}| \geq i \tag{22}
\]

**Proof:** The proof, based on Theorem 1 and Proposition 1, is omitted for the sake of brevity.

**Theorem 2 (Rule for Optimal Unconstrained Selection):** Let us consider a neighbor set \(\mathbb{N}\) with cardinality \(|\mathbb{N}| = n\) ordered according to the EOTX metric, namely \(f_{v_1} < f_{v_{i+1}}, \forall i < n\). The optimal unconstrained candidate set is one of the following sets:
\[
(v_1, v_2, \ldots, v_1, v_2, \ldots, v_n) \tag{23}
\]

**Proof:** The proof is a consequence of Propositions 1 and 2.

We note that, accounting for Theorem 2, the number of possible optimal unconstrained candidate sets is reduced from about \(n!\) to just \(n\).

**Remark (for Optimal Constrained Selection):** Let us consider a neighbor set \(\mathbb{N}\) with cardinality \(|\mathbb{N}| = n\) ordered according to the EOTX metric, namely \(f_{v_1} < f_{v_{i+1}}, \forall i < n\). The optimal constrained candidate set can be different from the optimal unconstrained candidate set, i.e. different from one of the following sets:
\[
(v_1, v_2, \ldots, v_1, v_2, \ldots, v_n) \tag{24}
\]

In fact, by assuming to limit the optimal candidate set to the size 1, the remark states that we can have:
\[
f(v_i) < f(v_1) \tag{25}
\]

Basing on (7) and after simple algebraic manipulation, we can write (25) as:
\[
f_{v_i} - f_{v_1} < \frac{d_i - d_1}{d_i d_1} \tag{26}
\]

which is for example true for \(f_{v_i} = 2, f_{v_1} = 1, d_i = 0.5, d_1 = 0.3\).

Accounting for Theorem 1 and the above remark, and assuming that the cardinality of the neighbor set is \(|\mathbb{N}| = n\),
the number of possible optimal \(m\)-constrained candidate sets is reduced from about \(n!\) to:

\[
\sum_{k=1}^{m} \binom{n}{k} = \sum_{k=1}^{n} \binom{n}{k} - \sum_{k=m+1}^{n} \binom{n}{k} = 2^n - 1 - \sum_{k=m+1}^{n} \binom{n}{k}
\]

(27)

where \(m \leq n\) and \(\binom{n}{k}\) is the number of unordered \(k\)-permutations of \(n\) elements without repetition.

B. Algorithm for Optimal Unconstrained Candidate Selection

Given the set of neighbors \(\mathbb{N}\) of a node \(s\), and for each neighbor \(v_i \in \mathbb{N}\) given the delivery ratio \(d_i\) and the minimum average number of opportunistic data-link transmissions \(f_{v_i}\), the Algorithm 1 singles out the optimum candidate set \(C_{\text{opt}}\) and the minimum average number of opportunistic data-link transmissions \(f_s\) according to Theorem 2, for which we have proved the optimality.

At the first, the algorithm sorts the neighbor set according to the EOTX metric, namely \(f_{v_i} < f_{v_{i+1}}, \forall i < n\) (see line 2). The optimal candidate set \(C_{\text{opt}}\) is initialized to the empty set and the minimum average number of opportunistic data-link transmissions \(f_s\) is set to infinite (lines 3-4). Then, each element of the neighbor set is incrementally included in the optimal candidate set (lines 6-12). We note that the inclusion of the new candidate in the candidate set must preserve the priority, i.e. the candidate is placed at the end of the candidate set (line 6). This process is repeated until the inclusion of a new candidate decreases \(f_s\) (lines 7-9). Otherwise, the optimum candidate set has been already discovered and therefore the procedure ends (lines 10-11).

The run-time complexity of the algorithm depends only on the size of the neighbor set \(n = |\mathbb{N}|\). In fact, the complexity of the sorting procedure is \(O(n \log n)\) elementary operations and the complexity of the for loop is \(O(n)\) aggregate operations, each of them consists of computing the probabilities \(p_i\) (3) and the EOTX \(f(C)\) (7). Since each aggregate operation can be executed in a constant time [9], the complexity of the whole algorithm is just \(O(n \log n)\), which is the same complexity exhibited by the algorithms for next hop selection in traditional shortest-path distance vector routing. Also the requirements of the algorithm in terms of information amount is the same of traditional distance vector routing: each node needs just to know the cost to reach each neighbor and the EOTX of the neighbor toward the other nodes in the network.

C. Algorithm for Optimal Constrained Candidate Selection

The Algorithm 2 singles out the optimum \(m\)-constrained candidate set \(C_{\text{opt}}\) and the minimum EOTX \(f_s\) for a candidate set of size \(m\) accounting for the Theorem 1.

The algorithm sorts the neighbor set according to the EOTX metric, namely \(f_{v_i} < f_{v_{i+1}}, \forall i < n\) (line 2) and then it builds the set \(C\) of the admissible candidate sets according to Theorem 1 (lines 3-8). More in detail, the set \(C\) is initialized to the empty set, and at each iteration of the first for loop the new candidate \(v_i\) is placed at the end of the already discovered candidate sets with size less than \(m\) (line 5-6). Then, the resulting candidate sets along with the set composed by the only node \(v_i\) are added to set of admissible candidate set (line 7). Once the admissible candidate sets with size less or equal to \(m\) are built, the algorithm looks for the optimal candidate set (lines 9-16).

The run-time complexity of the proposed algorithm is exponential with \(n\). In fact, the complexity of the sorting procedure is \(O(n \log n)\) elementary operations and the complexity of the construction of the admissible candidate sets is \(O(n)\) elementary operations. On the other hand, the complexity of looking the best candidate set is in the worst case \(O(|C|) = O(2^n)\) aggregate operations that can be executed in a constant time (see [9]). Therefore the run-time complexity of the whole algorithm in the worst case is \(O(2^n)\), which is considerably less than the factorial complexity of an exhaustive search on the whole set of candidate sets. As regards to the requirements of the algorithm in terms of information amount, the same considerations of Sec. IV-B hold.

Algorithm 1 Unconstrained Candidate Set Selection

1: // sorting \(\mathbb{N}\) in increasing order according to \(f_{v_i}\)
2: sort(\(\mathbb{N}\))
3: \(C_{\text{opt}} = \emptyset\)
4: \(f = \infty\)
5: for \(v_i \in \mathbb{N}\) do
6: \(\mathcal{C} \leftarrow \mathcal{C}_{\text{opt}} \cup \{v_i\}\)
7: if \(f(\mathcal{C}) < f\) then
8: \(\mathcal{C}_{\text{opt}} \leftarrow \mathcal{C}\)
9: \(f = f(\mathcal{C})\)
10: else
11: return \(\mathcal{C}_{\text{opt}}, f\)
12: end if
13: end for

Algorithm 2 \(m\)-Constrained Candidate Set Selection

1: // sorting \(\mathbb{N}\) in increasing order according to \(f_{v_i}\)
2: sort(\(\mathbb{N}\))
3: \(C = \emptyset\)
4: for \(v_i \in \mathbb{N}\) do
5: \(\mathcal{C}_{<m} \leftarrow \{\mathcal{C}\}, \forall \mathcal{C} \in \mathbb{C} : |\mathcal{C}| < m\)
6: \(\mathcal{C}_{\text{temp}} \leftarrow \mathcal{C}_{<m} \cup \{v_i\}\)
7: \(\mathcal{C} \leftarrow \{\mathcal{C}, \mathcal{C}_{\text{temp}}, (v_i)\}\)
8: end for
9: \(\mathcal{C}_{\text{opt}} = \emptyset\)
10: \(f = \infty\)
11: for \(\mathcal{C} \in \mathbb{C}\) do
12: if \(f(\mathcal{C}) < f\) then
13: \(\mathcal{C}_{\text{opt}} \leftarrow \mathcal{C}\)
14: \(f = f(\mathcal{C})\)
15: end if
16: end for
V. EXPERIMENTAL RESULTS

We validate the proposed algorithm for the constrained candidate-set selection using the link-level measurements of the MIT Roofnet topology [11]. The network of the test-bed is composed by 38 nodes 802.11b, and for each pair of nodes the delivery ratio has been estimated for different data rates: 1, 2, 5.5 and 11Mbps. Since the results for the different data rates exhibit similar behaviors, in the following we report only the experiments that refer to 1Mbps.

Fig. 1 presents a performance comparison of the proposed algorithm for 3-constrained selection with the algorithm for unconstrained selection in terms of minimum EOTX. Each dot $(x, y)$ in the figure represents the number of data link transmissions for the unconstrained $(x)$ and for the constrained algorithm $(y)$ for a given source-destination pair. The results clearly show that the performances in terms of opportunistic advantage are practically the same for both the algorithms. More in detail, the average (with respect to the pairs) minimum EOTX for the unconstrained selection is 2.4013, while the one for the 3-constrained selection is 2.4162 (2.4566 for $m = 2$ and 2.4064 for $m = 4$), i.e. for the considered test-bed the EOTX reduction is negligible for sets larger than 3 nodes.

Fig. 2 presents a comparison of the algorithms in terms of candidate set size. This experiment provides an insight about the cost of optimal candidate selection in terms of routing overhead. In fact, the larger is the selected candidate set, the larger is the amount of information to be delivered for candidate coordination. Therefore, the significant overhead required by the unconstrained candidate coordination is not justified since the performance gain is negligible.

Finally, Fig. 3 gives evidence of the non-optimality of the algorithm for unconstrained candidate selection when the size of the candidate set is limited (Sec. IV-A). As in Fig. 1, each dot in the figure represents the number of the data link transmissions for the 3-unconstrained and for the constrained algorithm, and we observe that the performances of the two algorithms differ significantly.

VI. CONCLUSION

In this paper we have proposed an analytical framework to model the problem of selecting the optimal candidate set for both the constrained and the unconstrained case. Moreover, we have proposed a distance vector algorithm for each case, and we have analytically proved their optimality. We underline that the algorithm for the constrained selection is the first solution able to reduce the complexity of the problem to an exponential factor. We validated the proposed algorithms with the link-level measurements of a 38-node 802.11 test-bed.

ACKNOWLEDGEMENT

The authors would like to thank Dr. C. Vincenzo Angelino and Dr. Pasquale Imperatore for their useful comments.

REFERENCES