Computer Networks 53 (2009) 25–44

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Mobility-aware clustering algorithms with interference constraints in wireless mesh networks

Rami Langar [a,b,*], Nizar Bouabdallah [c], Raouf Boutaba [a]

[a] School of Computer Science, University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada N2L 3G1
[b] LIP6, UMR 7606, University of Paris 6, 104 av. du President Kennedy, 75016 Paris, France
[c] INRIA-IRISA, Campus Universitaire de Beaulieu, 35042 Rennes Cedex, France

ARTICLE INFO

Article history:
Received 15 January 2008
Received in revised form 27 June 2008
Accepted 15 September 2008
Available online 1 October 2008

Responsible Editor: M.C. Vuran

Keywords:
Wireless mesh networks
Mobility management
Clustering
Radio resources
Interference
Path selection
Performance analysis

ABSTRACT

One of the major concerns in wireless mesh networks (WMNs) is the radio resource utilization efficiency, which can be enhanced by managing efficiently the mobility of users as well as the interference effect among neighboring links. To achieve this, we propose in this paper the use of clustering and a new interference-aware routing metric, called INX. Specifically, we first propose two mobility-aware clustering algorithms that take into consideration the mobility properties of users in order to improve the WMN performance. Then, we propose the use of INX in the clustering process in order to maximize the total network throughput. We prove through simulations that both clustering schemes can achieve significant gains in terms of radio resource utilization and load balancing, especially when using the INX metric. Hence, and as a main contribution, we show that by taking into account the interference effect between links, we can improve the performance of our clustering algorithms and increase the gain initially observed with the conventional hop-count metric.

1. Introduction

Wireless mesh networks (WMNs) provide the necessary ingredient needed for building bandwidth efficient and flexible networks to support the increasing demand for mobile wireless access to the Internet. A variety of applications are expected to benefit from WMNs such as "community wireless networks" [1–4]. Typically, a WMN is characterized by a two-tier architecture. The first layer comprises static wireless mesh routers, also called access points (APs). They represent the wireless backbone of the network. The second layer corresponds to the set of mobile users served by the APs and connected to the wired network through multi-hop wireless routing. Note that the mesh nodes connected directly to the wired network (i.e., connecting the WMN to the wired network) are called gateways. They represent the traffic sinks and sources to the WMN.

In WMNs, two types of messages are exchanged between the gateway and a mobile user: data and signaling messages. Each time the mobile user moves to a new AP, it notifies the gateway with its new location. These signaling messages are relayed through multi-hop wireless links to the gateway and compete for the scarce spectrum. As such, the amount of exchanged signaling messages is far from being negligible, especially for high mobility rate, and can account for a significant portion of the radio resources a mobile node uses in some cases. Hence, to achieve an efficient use of the radio resources in such multi-hop wireless networks, we need to minimize the average number of wireless links occupied by a mobile user during its service (i.e., while exchanging both data and signaling

* Corresponding author. Address: School of Computer Science, University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada N2L 3G1. Tel.: +1 519 7224312; fax: +1 519 8851208.

E-mail addresses: rami.langar@lip6.fr (R. Langar), nizar.bouabdallah@inria.fr (N. Bouabdallah), rboutaba@uwaterloo.ca (R. Boutaba).

1389-1286/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2008.09.012

packets). Henceforth, we referred to this metric as the radio resource utilization (RRU) cost.

Reducing the RRU cost per user improves the utilization efficiency of the WMN resources and thus increases the number of accepted subscribers to the WMN service. This means more profits to the service provider from the existing infrastructure. Note that, the RRU minimization must be achieved while ensuring the QoS requirements of users, such as the delay constraints.

One way to achieve this is by means of clustering as demonstrated in our study. Accordingly, the WMN is divided into a set of virtual clusters, covering all the nodes in the network. In each cluster, a node would serve as a clusterhead (CH). It operates as an intermediate node between the gateway and the APs inside the cluster. With regard to the management operations, the CH substitutes to the gateway inside the cluster and manages the mobility of local users. The main idea behind WMN clustering is to restrict a major part of the exchanged signaling messages due to the mobility of users to a local area (i.e., the cluster). As such, less wireless links are used by the signaling messages, which reduces the RRU cost and improves thus the network performance.

In this paper, we first propose two clustering schemes in order to minimize the RRU cost in WMNs. In the first scheme, called optimal static clustering (OSC), we assume that the clusters are static and disjoint. In this case, we determine the optimal cluster placement that minimizes the RRU cost by formulating the placement problem as an integer linear program (ILP). In the second scheme, called distributed clustering algorithm (DCA), the clusters may overlap. In this case, the cluster placement is done in a distributed manner. Typically, each AP calculates its own cluster when it acts as a CH. In essence, the DCA approach is proposed to alleviate the time complexity entailed by the OSC approach due to the time consuming resolution of the associated ILP problem. In this context, the cases of OSC and classical WMNs (i.e., without clustering) are used also to develop baselines to which the DCA improvements could be compared, since no clustering approach for mobility management in the context of WMNs has been previously proposed in the literature. It is worth noting that in both proposed clustering schemes, we take into consideration the mobility properties of the users as well as the interference effect among neighboring links. To the best of our knowledge, we are the first to introduce such constraints in the clustering process.

To mitigate interferences experienced by the wireless links, we also propose in this paper the use of a new routing metric, called interferer neighbors count (INX) during the clustering process. Indeed, the procedure of path selection fails to take into account the impact of such choice on the subsequent connection arrivals. Specifically, selecting the best current route for a connection in terms of bandwidth or loss rate may deteriorate severely the quality of the remaining available resources in the network due to the resulting interference. Consequently, new arriving connections will experience poor services. The philosophy behind our INX metric is then to choose a good path, not necessarily the best in terms of throughput, but that alleviates the resulting interference in order to preserve good paths for the subsequent arriving connections. As a result, the total network throughput will be maximized.

To gauge the effectiveness of our proposals, numerous simulations are performed. We first evaluate the potential performance gains of INX using ETX [5] and hop-count metrics as baseline examples. Then, we evaluate the performance of our clustering algorithms in arbitrary meshed topologies with different sizes and under various mobility and traffic scenarios using two routing metrics: the traditional hop-count and INX metrics. Our aim is to show the effectiveness of our algorithms when different WMN topologies and different routing metrics are used. We compare our proposals in terms of total network throughput, packet loss rate, and cumulative distributed function of delivery probabilities for used links, as well as registration updates cost, data delivery cost, RRU cost and load balancing.

The remainder of this paper is organized as follows. Section 2 presents an overview of the related works, followed by a description of the network model and the problem statement in Section 3. In Section 4, we introduce our new interference-aware routing metric (INX) used in the clustering process. A detailed description of the OSC method with its associated ILP formulation as well as the DCA algorithm are given in Section 5. Section 6 introduces the performance metrics used in our analysis. In Section 7, simulation results are provided. Finally, Section 8 concludes this paper.

2. Related work

Management of wireless multi-hop networks has been an active research area in the last few years. Numerous proposals for clustering schemes have been proposed in the literature in the context of ad-hoc networks. Clustering in such networks is introduced mainly to handle efficiently the frequent network topological changes due to ad-hoc nodes mobility [6–14].

In view of this [6] proposes a cluster-based framework, which defines a strategy for dynamically organizing the topology of an ad-hoc network due to the node mobility. This scheme is merely a cluster-based hybrid routing protocol, where the routes are discovered reactively for the inter-cluster routing and proactively inside the cluster. However, the interference effect between links has not been considered during the clustering process. In addition, no criteria are given to specify the optimal cluster organization.

Authors in [7] propose a mobility-aware routing protocol for mobile sensor networks using the concept of non-overlapping square zones or clusters. However, this scheme impose restrictions on the shape and the geographic location of clusters, by assuming all sensors in the network to be location-aware using GPS for example.

Likewise, the work in [8] proposes a mobility-resistant clustering approach for ad-hoc and sensor networks. However, this scheme takes into account the energy constraint of nodes while forming the clusters, which is not relevant in the context of WMNs.

So far, the above-mentioned clustering schemes have focused mainly on maintaining routes and adapting quickly to topological changes of the wireless backbone in ad-hoc networks, which occur only occasionally in WMNs due to AP failures or AP addition. Indeed, in WMNs, it is the nodes located at the second level of the WMN architecture (i.e., the users) that are mobile. Whereas, the nodes located at the first level of the WMN architecture (i.e., APs) and forming the wireless backbone, are static. In our work, we try to convey information regarding the mobility of users to the upper layer of WMNs formed by static APs in order to optimize the WMN performance. Hence, designed for ad-hoc networks, the aforesaid protocols are not suitable for typical WMN applications for two main reasons: the static topologies of WMNs and the different communication patterns. Specifically, in WMNs, almost all traffic is either to or from a gateway, while in ad-hoc networks, the traffic flows between arbitrary pairs of nodes.

In [15], the author presents a clustering algorithm for "quasi-static" ad-hoc networks, where nodes are static or moving at a very low speed. The proposed scheme is more adapted to the WMN environment. However, it is concerned with 1-hop clustering, which defeats the purpose of clustering in WMNs. Indeed, dividing the WMN into small-size clusters (i.e., with a radius equal to 1) will result in an increasing number of the expensive inter-cluster handoffs instead of the low-cost local intra-cluster handoffs during user mobility. In addition, the interference effect has not been considered during the clustering process.

Other works [16,17] proposed *d*-hop clustering algorithms in wireless ad-hoc networks such that a node in any cluster is at most *d*-hops away from the CH. The proposed schemes are not restricted to 1-hop clustering. However, the clusters have the same radius *d*, an additional constraint, which may lead to unsatisfactory results regarding the RRU cost minimization. Again, because of the unique characteristics of WMNs, existing ad-hoc clustering approaches can not be directly used.

Another clustering approach in mobile ad-hoc networks, based on graph theory, namely connected dominating set (CDS), is used in [18–20]. In this approach, the objective is to identify the smallest set of CHs that forms a CDS. This problem is known in graph theoretic terminology as minimum connected dominating set (MCDS). Hence, by definition of a CDS, each ad-hoc node not included in the set of CHs has at least one adjacent node belonging to the CH set. The set of CHs operates therefore as routers and forms a virtual backbone for the ad-hoc network. Again, it is easy to see that the proposed scheme is concerned with 1-hop clustering, which defeats the purpose of WMNs.

In the context of WMNs, relevant works to clustering are [21–24]. These works attempt to integrate the WMNs with the wired backbone. Typically [21–24], investigated the well known problem of gateway placement in WMNs. The problem consists in dividing the WMN into a minimum number of disjoint macro-clusters, where each macro-cluster is assigned to a gateway node that connects directly to the wired network. The objective is therefore to minimize the number of deployed expensive gateways (i.e., number of macro-clusters) required to connect all APs to the wired network subject to several QoS constraints such as the gateway capacity, the cluster radius, etc. In our study, we focus rather on virtual clustering inside each macro-cluster, while taking into consideration the mobility properties of users as well as the interference effect between links. To the best of our knowledge, we are the first to introduce such constraints in the clustering process. Our objective is to divide the macro-cluster into virtual micro-clusters in order to minimize the RRU cost in the WMN, and then maximize the network throughput. Hence, while macro-clustering is performed to minimize the number of required gateways, micro-clustering is performed to optimize the RRU cost in WMNs. Both clustering approaches are complementary to achieve cost-effective WMNs.

3. System model and problem description

3.1. System model

We represent a wireless mesh network (WMN) by a directed graph $G(V,E)$, called a connectivity graph G. Each node $v \in V = \{1,\ldots,N\}$ represents an Access Point (AP), which is characterized by a circular transmission range $R_t(v)$ and a carrier sensing range $R_h(v)$ (called also hearing range). During the transmission of a node υ, all nodes inside its carrier sensing range, denoted by $H(v)$, sense the channel to be busy and cannot access the medium. Hereafter, we denote by $H^+(v) = H(v) \cup \{v\}$ and by $H^-(v)$ the set of nodes that node υ cannot hear, i.e., $H^-(v) = V \setminus H^+(v)$.

On the other hand, during the transmission of the node υ, all the nodes residing in its transmission range, and thus representing its neighborhood denoted by $N_e(v)$, receive the signal from υ with a power strength such that correct decoding is possible with high probability. A bidirectional wireless link exists between υ and every neighbor $u \in N_e(v)$ and is represented by the directed edges (u,v) and $(v,u) \in E$. The number of neighbors of a vertex υ is called the connectivity degree of υ, denoted by $\delta(v)$. The average connectivity degree $\bar{\delta}$ of a graph G is called the graph degree and defined by $\frac{1}{|V|}\sum_{v\in V}\delta(v)$.

We represent the graph connectivity by a connectivity matrix. The connectivity matrix of $G(V,E)$ is a matrix with rows and columns labeled by the graph vertices V, with a 1 or 0 in position (v,u) according to whether υ and u are directly connected or not.

Each node υ is equipped with a 802.11 wireless card, and transmits with a rate $r_{v,u}$ to its neighbor $u \in N_e(v)$. The available transmission bit rate $r_{v,u}$ is determined by the link condition. Typically, $r_{v,u}$ depends on the distance between nodes υ and u and the resulting loss rate on that link. Different algorithms exist to choose among the available transmit bit rates. For example *SampleRate* [25,26] selects the bit rate based on the delivery probabilities measured at the different bit rates. A simpler alternative consists in fixing the transmission bit rates to a given value as in [27], where the transmission bit rate of the 802.11 radios was limited to its lowest sending rate of 1 Mbit/s.

In our system model, all the radios share the same channel. The RTS/CTS mechanism is disabled. It was shown in [25,28] that RTS/CTS increases the overhead without really improving the network performance. In view of this, the basic access mode (i.e., Data/ACK) is considered.

It is worth noting that once a node υ accesses successfully to the shared medium, it blocks the transmission of all nodes inside its carrier sensing range (i.e., $H(v)$) during the transmission time of the data packet, denoted by T_{data}. The transmission of the node υ may also prevent the nodes in $H(v)$ to receive correctly packets from their neighbors. On the other hand, a successful transmission from υ to u requires all the hidden nodes from υ (i.e., $H^{-}(v) \cap H(v)$) to be silent during $T_{\text{data}} + \text{SIFS}$ (called also vulnerable period [29]), otherwise the receiving node u will not return an ACK frame to υ. If this happens, node υ considers the packet as unsuccessfully transmitted and schedules for a retransmission later. Consequently, a successful packet delivery from υ to u needs or results in the silence of each wireless link (m,n) (i.e., $d(m,n) \leqslant R_t(m)$) verifying the following condition:

$$\begin{cases} D(v,m) \leqslant R_h(v) \text{ or } D(v,n) \leqslant R_h(v) \text{ or } D(m,u) \leqslant R_h(m): \\ \text{during the transmission of the data packet from } v \text{ to } u, \\ D(u,m) \leqslant R_h(u) \text{ or } D(u,n) \leqslant R_h(u) \text{ or } D(m,v) \leqslant R_h(m): \\ \text{during the transmission of the ACK frame from } u \text{ to } v \end{cases} \tag{1}$$

where $D(m,n)$ denotes the Euclidian distance between nodes m and n. If the link (m,n) verifies the condition (1), it is referred to as interferer link to the link (v,u). We call $S(v,u)$ the set of interferer links of the link (v,u). Note that in (1), we have not considered the signal capture property.

Based on the interferer link definition, we associate to the connectivity graph G a conflict matrix C. The conflict matrix C of $G(V,E)$ is a matrix with rows and columns labeled by the vertices V, such that in position (v,u):

$$C_{vu} = \begin{cases} |S(v,u)| & \text{if nodes } v \text{ and } u \text{ are directly connected}, \\ 0 & \text{otherwise} \end{cases}$$

where $|S(v,u)|$ represents the number of interferer links resulting from the use of the link (v,u). Note that $C_{vu} = 0$ means that υ and u are not directly connected; otherwise the graph G is not connected.

The number of interferer links of a path is defined as the sum of the number of interferer links of each link in the path. We denote by I the interference matrix, which represents the number of interferer links of the paths connecting the mesh nodes. Thus, I_{vu} denotes the number of interferer links of the path used between pair of nodes (v,u). Clearly, the matrix I depends on the used routing policy (i.e., the adopted routing metric).

3.2. Problem description

In this paper, we focus on optimizing the radio resource utilization (RRU) cost of a mobile user during its service, and at the same time taking into consideration the interference effect among the neighboring links. The RRU cost includes that of exchanged data packets as well as of the signaling messages used to manage the user mobility. The aim is to improve the utilization efficiency of the WMN resources, which increases the number of accepted connections in the access network, resulting thus in more profit for service providers.

Note that the RRU cost minimization is achieved while ensuring the QoS requirements of users, such as the delay constraints. In our work, we define $I_{\max}$ as the maximal tolerable interference allowed inside the WMN. In other words, $\forall u, v \in V$, $I_{uv} \leqslant I_{\max}$. Such constraints guarantee that the connection between u and υ satisfies certain delay and bandwidth requirements.

To achieve this, we use clustering. Three hierarchical levels in the WMN are identified in this case. At the top of the hierarchy is the gateway node, which is connected directly to the wired network for Internet access and serves all the mesh nodes (i.e., APs) in the WMN. The second level of hierarchy is the CH, serving the APs inside the cluster, and the third level is the AP offering IP-layer connectivity to mobile users located within its transmission range (i.e., inside the AP subnet).

For mobility management, we define two types of handoffs: Intra-cluster and Inter-cluster handoffs. An Intra-cluster handoff occurs when a user moves between two APs that belong to the same cluster. On the other hand, an Inter-cluster handoff occurs when a user moves between two APs belonging to different clusters. To maintain connectivity during user mobility, the gateway keeps a record of the current user cluster (i.e., the current CH's identity). Each time the user crosses a cluster boundary (i.e., Inter-cluster handoff), it updates the system with its new location by sending a registration update message to the gateway through the CH of the new visited cluster. We call this kind of registration updates GW registration. In contrast, when an Intra-cluster handoff occurs, the update registration message will be sent only to the current CH and will not be forwarded to the gateway. This kind of registration is called CH registration.

With regard to data packets, an incoming packet from the backbone to the mobile user (i.e., downlink traffic) is first intercepted by the gateway. Then the packet is forwarded to the current user CH, which relays the data packet to the corresponding AP for delivery. In turn, data packets transmitted by the mobile user to the wired network (i.e., uplink traffic) are directly routed to the gateway, for instance, according to the shortest path, without requiring to pass-through the CH.

The RRU cost of a mobile user involves two terms, i.e., the first one regarding the data packets' resource utilization and the second term is related to the resource utilization of the signaling messages used to manage the user mobility. We refer to the first term as the *data delivery cost* and to the second term as the *registration updates cost*.

Clearly, an efficient clustering policy must minimize the sum of these two terms. It must achieve a balance regarding the cluster sizes in terms of number of APs. Specifically, a WMN with small-size clusters will result in an increasing number of expensive inter-cluster handoffs instead of low-cost local intra-cluster handoffs during the user mobility. As such, the registration updates cost increases. On the other hand, large-size clusters will reduce the number of expensive inter-cluster handoffs. However, this will result in an increase of the data delivery cost since more links will be involved in the routing process between the gateway and the mobile user through the CH. In view of this, the registration updates cost and the packet delivery cost are two opposite requirements. A tradeoff between these two requirements must be achieved by optimally constructing the clusters in order to minimize the RRU cost.

To achieve the above-mentioned tradeoff, we suggest two clustering schemes. We first assume that the clusters are static and disjoint. In this case, we determine the optimal cluster placement that minimizes the RRU cost by formulating the placement problem as an integer linear program (ILP). As an alternative to the static clustering, we propose a distributed algorithm to construct the virtual clusters. In this case, the clusters are neither static nor disjoint and their placement depends on the user mobility (i.e., trajectory). This method is introduced to alleviate the long computing time required to resolve the static clustering ILP problem. A detailed description of both clustering schemes is given in Section 5.

It is worth noting that the path selection during the clustering process depends mainly on the used routing policy. Hence, to achieve further gain in terms of RRU cost, an interference-aware routing metric must be used in the clustering algorithm. To do so, we propose a new routing metric, called interferer neighbors count (INX). Indeed, selecting the best current route for a connection in terms of bandwidth or loss rate may deteriorate severely the quality of the remaining available resources in the network due to the resulting interference. Consequently, new arriving connections will experience poor services. Hence, the philosophy behind our INX metric is to choose a good path, not necessarily the best in terms of throughput, but that alleviates the resulting interference in order to preserve good paths for the subsequent arriving connections. As a result, the total network throughput will be maximized.

In the remainder of this paper, we will first evaluate our proposed INX routing metric using ETX [5] and hop-count as baselines. Then, we compare between the following networks based on their RRU costs: classical WMN (i.e., without clustering (WC)) and cluster-enabled networks with both static and distributed cluster placement strategies, using two routing metrics: hop-count and INX. The aim is to study the impact of interference on our clustering algorithms. Note that the classical and static cluster-enabled cases are used as baselines to which the DCA improvements could be compared.

4. The interference-aware routing metric: INX

This section introduces our new routing metric, called interferer neighbors count (INX) that takes into consideration the number of interferer links, defined in Section 3.1. Recall that this metric will be used in a latter stage for the clustering process in order to study the impact of interference on our clustering algorithms. In addition, to gauge the efficiency of our metric, it will be compared with its peers, such as ETX [5], and with the traditional hop-count metric.

4.1. Motivating the INX routing through an example

To achieve efficient spectral utilization and high throughput in multi-hop networks, the path selection within the mesh network is crucial. To this end, several routing metrics have been proposed in the literature [5,32,33]. Hop-count, which is the most commonly used routing metric in existing routing protocols such as ad-hoc on-demand distance vector (AODV [34]) and destination sequenced distance vector (DSDV [35]) treats all links in the network to be alike and finds paths with the shortest number of hops. However, it does not account for interference experienced by the wireless links, which can often result in paths with high loss ratio and poor performance [5,32].

To overcome these limitations, quality-aware routing metrics [5,32,33] have been proposed to improve the throughput of WMNs. For instance, the expected transmission count (ETX) [5] specifies the quality of a link in terms loss ratio. It is defined as the expected number of MAC retransmissions needed to successfully deliver a packet from the sender to the receiver. The lower the ETX metric for a link, the better is the link. The path metric is the summation of ETX of each link in the path.

Although the performance improvement of ETX over le minimum hop-count metric is impressive, it does not cope well with the inter-path interference when there are multiple flows in the network. A successful transmission over a link $l = (v, u)$ either needs the silence or prevent the transmission over the interferer links of l. Specifically, a first subset of source nodes in the WMN (i.e., $H(v)$) will be automatically *blocked* since the common channel is sensed as busy during the transmission of υ. A second subset of receiving nodes will not be able to receive correctly from their neighbors (i.e., due to collisions) since they receive at the same time the in-progress transmission of υ. Finally, a third subset of nodes (i.e., hidden nodes from v) have to remain silent to avoid collisions at u with the transmitted packet by υ.

To allow the maximum simultaneous transmission within the WMN, and then improve the total network throughput, the blocked interferer links need to be minimized. However, the above routing metric pay little attention to employ such information to build paths. To illustrate this, let us consider the six-node WMN of Fig. 1, where each wireless link is associated with its ETX metric.

Assume that two connections requests C_1 and C_2 arrive successively. The first arriving connection C_1 that needs to be served is from A to D. According to the ETX-based routing, this connection is carried through the path $\mathscr{P}_1 = [A, B, D]$ instead of $\mathscr{P}_2 = [A, C, D]$ since $\mathscr{P}_1$ achieves better ETX value (i.e., 2.8 < 3.2). Meanwhile, the second connection request C_2 from E to F arrives. This connection will be blocked until the service of the current packet is finished since node E is already busy due to the transmission of the intermediate node B on $\mathscr{P}_1$. In other words, the link (E, F) is an interferer link with regard to the link $(B, D) \in \mathscr{P}_1$. We can see that connection C_1 and C_2 interfere and the service of one of them result in the blockage of the second one. Hence, the total throughput is divided by a factor of two.

Alternatively, if C_1 was served through the path $\mathscr{P}_2$, the connection C_2 would be satisfied through the link (E, F) without any interference with the C_1. Hence, both connections can be served simultaneously, in the sense that nodes

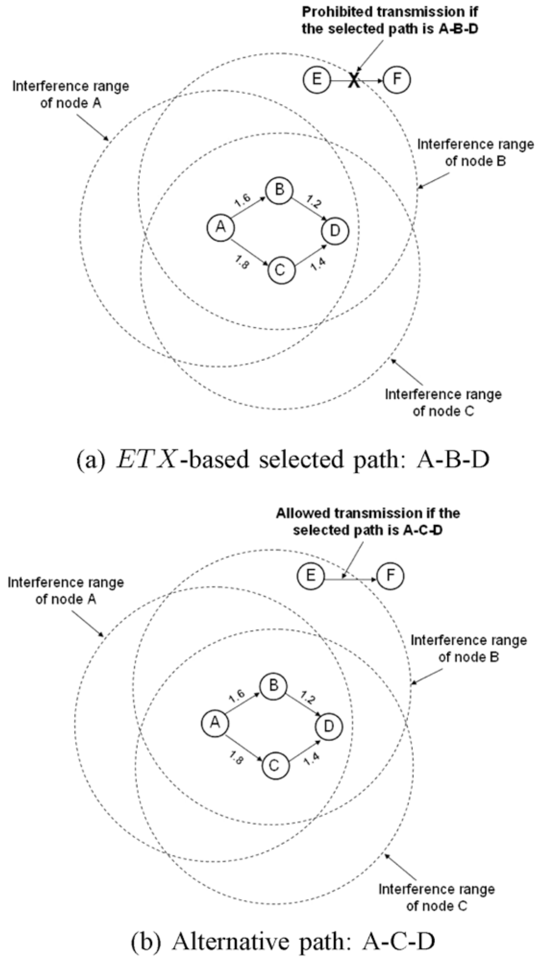


Fig. 1. Understanding inter-path interference.

D and F can receive packets at the same time from their respective sender nodes A and E.

This example shows the limitation of the ETX metric to learn enough about its environment in order to maximize the number of accepted connections. It is worth noting that the ETX metric may select the best path with the highest throughput for a given connection. But this choice does not necessarily ensure the highest total network throughput. Typically, the ETX-based routing enables to find the best current route for an arriving connection, however, it does not consider the impact of that choice on the resulting network state.

4.2. The INX routing metric

In this subsection, we propose to extend the ETX metric to take into account the interference experienced by the wireless links. The new metric is called interferer neighbors count (INX). The INX of a link (v,u) is defined as the product of the ETX of the link (v,u) and the number of all the interferer links resulting from a transmission on that link (v,u) (i.e., the parameter C_{vu} of the conflict matrix C) weighted by their respective bit rates as follows:

$$\mathrm{INX}(v,u) = \mathrm{ETX}(v,u) \sum_{\text{link } (m,n)\in S(v,u)} r_{m,n}. \quad (2)$$

Clearly, The INX metric captures the quality of the wireless link by including the ETX metric. We also want to consider the impact of using this link on subsequent connection arrivals. Simply counting the number of interferer links will not ensure correctly this property, since we are distinguishing between good (high throughput) and bad (low throughput) interferer links. To reflect this, our metric includes the bit rate of each interferer link.

For illustration, let revisit the example of Fig. 1. Assume that all the wireless links have the same bit rate 1 Mbit/s. $C_{AB} = |S(A,B)| = |\{BA, AC, CA, CD, DC, BD, DB, EF\}| = 8$, hence $\mathrm{INX}(A,B) = 1.6 \times 8 = 12.8$. On the other hand $\mathrm{INX}(A,C) = 1.8 \times 7 = 12.6$.

The path metric $\mathrm{INX}(\mathscr{P})$ for a given path $\mathscr{P}$ is the normalized summation of the INX of each link in the path and can be written as follows:

$$\mathrm{INX}(\mathscr{P}) = \frac{1}{N_l} \sum_{\text{link } l\in\mathscr{P}} \mathrm{INX}(l), \quad (3)$$

where N_l is the total number of directional links in the WMN.

Note that our proposed metric can be easily implemented and combined with any existing routing metric (e.g., ETT [33], WCETT [33]). Moreover, if two or more paths experience the same value of INX, the one with the minimum number of hops will be selected. Revisiting again the example of Fig. 1. We have $\mathrm{INX}(\mathscr{P}_1) = \frac{1}{10} \times (1.6 \times 8 + 1.2 \times 9) = 2.36 > \mathrm{INX}(\mathscr{P}_2) = \frac{1}{10} \times (1.8 \times 7 + 1.4 \times 7) = 2.24$. As such, the path $\mathscr{P}_2$ will be used to carry the connection C_1 instead of $\mathscr{P}_1$ as dictated by the basic ETX-based routing. In doing so, the arriving connection C_2 will start its service without waiting for the end of the service of C_1.

In what follows, we present our mobility-aware clustering algorithms in WMNs that use the above metric in the path selection during the clustering process.

5. Mobility-aware clustering algorithms

As stated before, we propose two mobility-aware clustering algorithms that aim at minimizing the RRU cost in WMNs, while taking into account both the mobility properties of users and the interference effect between links.

5.1. Optimal static clustering

In this approach, we address the optimal placement of clusters in WMNs by deriving a theoretical lower bound based on linear programming. We aim at minimizing the RRU cost of a mobile user subject to its QoS requirements in terms of delay or interference constraints.

Recall that in our work, we defined $I_{\max}$ as the maximal tolerable interference allowed inside the WMN. In other words, $\forall u, v \in V$, $I_{uv} \leqslant I_{\max}$. Such constraints guarantee that the connection between u and v satisfies certain delay and bandwidth requirements.

We formulate the clustering problem as follows. Given a WMN of N nodes, find the disjoint sets of APs (i.e., clusters) that minimizes the total radio resource utilization subject to the QoS constraints. Note that the disjoint clusters must cover all the nodes of the network and satisfy the connectivity constraint. That is, each cluster must be connected, in the sense that starting from any node of the cluster, one should be able to reach all other nodes of the cluster without leaving the cluster. In the following, we give an ILP formulation of the above problem.

5.2. ILP formulation

Let $N = |V|$ be the number of APs. The APs are denoted by AP_i, $(i = 1,\ldots,N)$. We denote by $AP_1 \in V$ the gateway (GW) that connects the WMN to the wired network. We introduce a binary variable a_i to indicate whether an $AP_i \in V$ is set up as a CH or not. To represent CHs allocation for APs, we define another binary variable $b_{i,j}$ which takes the value of 1 whenever AP_j, $j = 1,\ldots,N$, is assigned to the CH AP_i, $i = 1,\ldots,N$. M is an upper bound on the number of clusters that can be formed. m_{sig} and m_{data} represent the average size of signaling messages used for registration updates and the average size of data packets, respectively. $1/\mu$ represents the mean sojourn time of a mobile user in a subnet (i.e., AP), and λ is the downlink packet transmission rate (in terms of packets/s). For each user inside the WMN, let also its mobility pattern be defined by the process $\{Y(t), t \geqslant 0\}$, where $Y(t) = AP_i$ represents the user's location at time t.

Given the matrix of interference cost I between APs, the probability Π_i that the mobile user is located at the subnet AP_i, and the transition probability matrix P of the process Y, our objective function will be to minimize the RRU cost in the WMN. It is worth noting that the interference cost matrix I depends on either the hop-count or the INX routing metric is used for path selection. In the former case, $I(u,v) = I_{\text{hop}}(u,v)$ corresponds to the interference level resulting from a transmission of a packet along the shortest path between nodes u and v. However in the latter case, $I(u,v) = I_{\text{interf}}(u,v)$ corresponds rather to the minimum interference cost (in terms of number of interferer links) generated by a transmission of a packet between these two nodes. Note that in both cases, I_{hop} and I_{interf} are symmetric matrices.

The RRU cost can be expressed as follows:

$$RRU\_Cost = \alpha \times Reg\_Update\_Cost + \beta \times Data\_Delivery\_Cost, \tag{4}$$

where $\alpha = \frac{2\,\mu m_{\text{sig}}}{2\,\mu m_{\text{sig}} + \lambda m_{\text{data}}}$ and $\beta = \frac{\lambda m_{\text{data}}}{2\,\mu m_{\text{sig}} + \lambda m_{\text{data}}}$ are the proportion of the amount of signaling messages and the proportion of data packets among the total traffic generated by a mobile user, respectively. The *Reg_Update_Cost* can be written as follows:

$$Reg\_Update\_Cost = \frac{1}{N_l} \times \sum_{i=1}^{N} \Pi_i \times Update\_Cost(i), \tag{5}$$

where N_l is the total number of directional links in the WMN, and $Update\_Cost(i)$ is the interference cost (in terms of number of blocked links) of the registration updates when the mobile user leaves the AP_i. It is given by:

$$Update\_Cost(i) = \sum_{j=1}^{N} (P(i,j) \times \sum_{k=1}^{N} b_{k,j} \times [I(j,k) + (b_{k,j} - b_{k,i}) I(k, GW)]), \tag{6}$$

where $I(i,j) = I_{\text{interf}}(AP_i, AP_j)$ or $I_{\text{hop}}(AP_i, AP_j)$ according to the used routing policy, and $P(i,j) = P(AP_i, AP_j)$ denotes the transition probability from AP_i to AP_j.

Likewise, the *Data_Delivery_Cost* can be expressed as follows:

$$Data\_Delivery\_Cost = \frac{1}{N_l} \times \sum_{i=1}^{N} \Pi_i \times Delivery\_Cost(i), \tag{7}$$

where $Delivery\_Cost(i)$ is the data delivery cost of downlink traffic when the mobile user is connected to AP_i. It is given by:

$$Delivery\_Cost(i) = \sum_{j=1}^{N} b_{j,i} \times [I(GW, j) + I(j, i)]. \tag{8}$$

Note that in our study, we consider only the packet delivery cost of the downlink traffic. This is because the packet delivery cost in the uplink direction (i.e., from the mobile user to the gateway) is the same for all the approaches (i.e., distributed clustering, static clustering and without clustering) since the path with minimum interference level (respectively minimum number of hops) is always used in this direction if the interference-aware routing metric (respectively the hop-count routing metric) is used for path selection. So, we omit this term as it is the same for all the compared approaches.

Hence, our ILP problem can be formulated as follows with the objective function:

$$\min \quad \{RRU\_Cost\}$$

subject to:

$$\begin{aligned}
&(a)\quad \forall j = 1,\ldots,N : \sum_{i=1}^{N} b_{i,j} = 1\\
&(b)\quad \forall i,j = 1,\ldots,N : a_i \geqslant b_{i,j}\\
&(c)\quad a_1 = a_{GW} = 1\\
&(d)\quad \forall i = 1,\ldots,N : a_i = b_{i,i}\\
&(e)\quad \sum_{i=1}^{N} a_i \leqslant M\\
&(f)\quad \forall j = 1,\ldots,N : \sum_{i=1}^{N} b_{i,j} \times (I(GW,i) + I(i,j)) \leqslant I_{\max}\\
&(g)\quad \forall i = 1,\ldots,N : a_i \in \{0,1\}\\
&(h)\quad \forall i,j = 1,\ldots,N : b_{i,j} \in \{0,1\}.
\end{aligned}$$

Condition (a) denotes that each AP is assigned to one and only one CH. Inequality (b) implies that a CH has to be set up before being assigned APs. Inequalities (c) and (d) ensure that the gateway can not be assigned to another cluster and each CH belongs to the cluster that it manages. Inequality (e) provides an upper bound on the number of the constructed clusters that can be parameterized by the WMN administrator. For instance, assigning to M the value of 1, implies a WMN without clustering. On the other hand putting $M = N$, implies that all the APs have the capability to operate as CHs. Inequality (f) traduces the interference constraint. The last two conditions indicate that a_i and $b_{i,j}$ are binary variables. Note that in the case of hop-count based routing, the inequality (f) will be replaced by

$$(f') \ \forall j = 1, \ldots, N : \sum_{i=1}^{N} b_{i,j} \times (d(GW, i) + d(i,j)) \leqslant D_{\max},$$

where $D_{\max} = \mathrm{Max}(\mathrm{Min}(d(GW, AP_i)))$, $i = 1, \ldots, N$ and $d(i,j) = d(AP_i, AP_j)$ is the distance between AP_i and AP_j.

We will show in Section 6, how to derive the vector Π and the transition probability matrix P, which are used as input to our ILP formulation.

In practice, once the clusters are identified by solving the ILP problem, a *Cluster Table* at each CH is implemented. The table associated to a given CH contains the set of APs assigned to that CH. Each AP retains also the identity of its corresponding CH. Finally, the mobile user maintains the identity of the CH of its connecting AP during its movement. Once, the mobile user moves to a new AP, it registers by sending a signaling message to the new AP containing the identity of its current CH. This registration message will be forwarded by the new AP to its CH. Accordingly, the CH achieves either a local registration (i.e., only a *CH registration*) or a *GW registration* by forwarding the received message to the gateway. Specifically, the CH compares its identity to the current CH identity of the mobile user. If the mobile user remains in the same cluster (i.e., the old and the new CH are the same) a simple local registration is achieved, otherwise a *GW registration* is performed.

5.3. Distributed clustering algorithm

Motivated by the dynamic and distributed nature of the clustering protocol operations, we now propose a distributed clustering policy, called DCA, that aims at minimizing the RRU cost subject to the delay or interference constraints. This distributed policy is a new alternative to divide the WMN into clusters while avoiding the time complexity of the static approach based on the time consuming ILP problem resolution.

Like the OSC approach, the DCA algorithm divides the WMN into virtual clusters where the mobile user limits its registration updates within this local area, instead of communicating to the far away gateway. The mobile user keeps registering with the current CH instead of the gateway as long as it moves inside the current virtual cluster. As a distinguishing key feature, the DCA clusters are constructed in a distributed manner and may overlap as opposed to the disjoint and centrally calculated OSC clusters. Moreover, the virtual cluster construction with DCA depends on the mobile user trajectory. It depends on the relative position of the old and the new APs with respect to the gateway and on the delay or interference constraint.

Specifically, according to the DCA approach, each time the mobile user moves to a new AP, it first compares the data delivery cost of its indirect path to the gateway through the current CH with a certain threshold *Thresh*. We refer to this operation as the data delivery cost verification. If this cost is equal or less than *Thresh*, the mobile user can register locally to the current CH (i.e., *CH registration*). Otherwise, it registers directly to the gateway (i.e., *GW registration*) and the new AP becomes the CH of the new cluster. As such, the new AP is considered as outside the previous cluster. Moreover, to minimize the signaling cost (i.e., registration updates cost), a second condition regarding the registration updates cost must be verified before performing a *CH registration* instead of a *GW registration*. Specifically, a *CH registration* is achieved as long as it is cheaper than a *GW registration*. Indeed, each time the mobile user moves to a new AP, the latter compares the signaling cost of a registration update to the CH with that to the gateway. If it is equal or higher than that to the gateway, a *GW registration* is preferred. Thus, the new AP is considered as not belonging to the previous cluster.

It is worth noting that the above data delivery and signaling costs verifications hardly depend on the used routing policy (i.e., the adopted routing metric used for path selection). Indeed, in the case of using the traditional hop-count routing metric, these two conditions can be written as follows.

AP_i belongs to the cluster managed by AP_j if and only if

$$\begin{cases} d(AP_i, AP_j) < d(AP_i, GW) \\ \& \\ d(AP_i, AP_j) + d(AP_j, GW) \leqslant D_{\max} \end{cases}, \tag{9}$$

where $d(x, y)$ denotes the shortest path distance (in terms of number of hops) between x and y, and $D_{\max}$ the maximum tolerable delay inside the WMN, expressed in terms of hops.

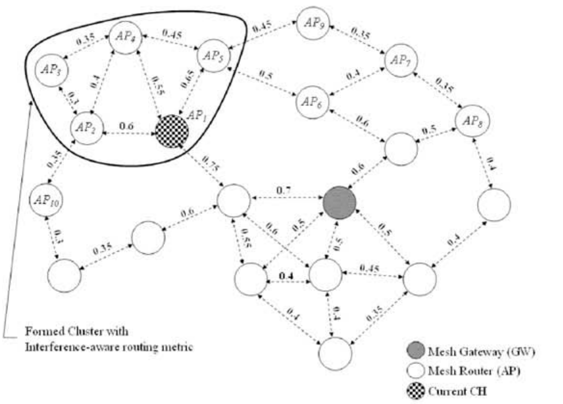
On the other hand, when using our INX routing metric, the above two conditions can be written as follows.

AP_i belongs to the cluster managed by AP_j if and only if

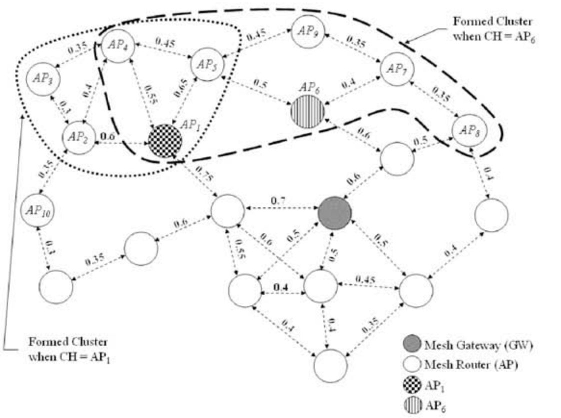
$$\begin{cases} \mathrm{INX}(AP_i, AP_j) < \mathrm{INX}(AP_i, GW) \\ \& \\ \mathrm{INX}(GW, AP_j) + \mathrm{INX}(AP_j, AP_i) \leqslant I_{\max} \end{cases}. \tag{10}$$

To illustrate the DCA algorithm, let us consider the two simple examples presented in Figs. 2 and 3 of a 20-node arbitrary wireless mesh topology when using two routing policies: the INX and the hop-count metrics. The WMN contains one gateway subrounded by 19 subnets (i.e., APs) randomly distributed. In these figures, we present the INX of each wireless link. The threshold $I_{\max}$ is defined as $\mathrm{Max}(\mathrm{Min}(\mathrm{INX}(GW, AP_i)))$, $i = 1, \ldots, 20$ and it is equal to 2.35 in our examples.

Assume that the user activates his mobile terminal at AP_1 as shown in Fig. 2a. Instantaneously, the mobile terminal registers to the gateway through the AP_1, which becomes the mobile user CH. The associated cluster to the current CH (i.e., AP_1) is composed of 5 APs as shown in this figure. These APs satisfy both conditions regarding the interference costs shown in (10). Accordingly, as long as the mobile user remains in this cluster (i.e., it fulfills the data delivery constraint and a *CH registration* is cheaper than a *GW registration*), it carries out a local registration with the CH. Once it leaves this cluster, it performs a *GW registration* and the new serving AP becomes the new CH of the new cluster.



(a) Formed Cluster when $CH = AP_1$ and $I_{max} = 2.35$



(b) Overlapping cluster property

Fig. 2. The DCA clustering scheme: an example of cluster placement in an arbitrary wireless mesh network using the Interference-aware routing metric, INX.

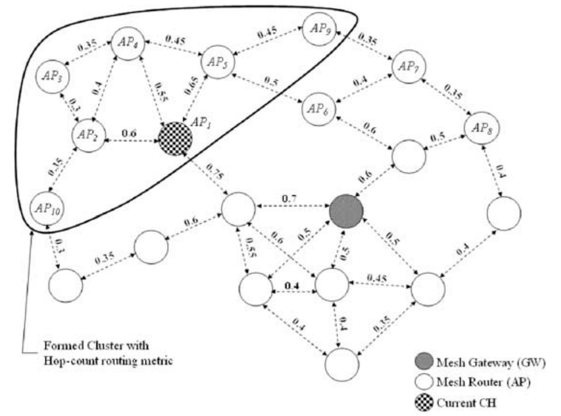


Fig. 3. The DCA clustering scheme: an example of cluster placement in an arbitrary wireless mesh network using the hop-count routing metric.

Specifically, assume that the mobile user moves from AP_1 to the neighboring subnet AP_6 along the following trajectory $\{AP_1, AP_2, AP_3, AP_4, AP_5, AP_6\}$ as depicted in Fig. 2b. The DCA algorithm operation can be described as follows. The mobile user begins its trajectory at subnet AP_1, which is set up as the current CH. When the mobile user moves to AP_2, it compares first the signaling cost between the local and gateway registrations. In this case, a *CH registration* cost (i.e., 0.6) is lower than that of the *GW registration* (i.e., 2.05). In addition, the data delivery constraint is respected as the cost of the indirect path from the gateway to AP_2 through the current CH AP_1 does not exceed the threshold I_{max}. Consequently, the mobile user performs a local registration (i.e., *CH registration*) and AP_2 joins the cluster managed by AP_1. Likewise, the mobile user achieves always a local registration to the CH (i.e., AP_1) when it moves to AP_3, AP_4 and AP_5 since the two above conditions are satisfied. As a result, these latter APs join also the cluster managed by AP_1.

The condition on the data delivery constraint is no longer satisfied when the mobile user enters subnet AP_6. Hence, the mobile user registers directly to the gateway and the new visited subnet AP_6 becomes the new CH of the new cluster. In this case, the new cluster managed by AP_6 will be composed of seven APs as shown in Fig. 2b.

We can see that AP_1, AP_4 and AP_5 belong to both clusters managed by CHs AP_1 and AP_6. So, in contrast to the OSC approach, the clusters are no longer disjoint and my overlap. Indeed with DCA, the mobile user can be attached to different CHs when visiting the AP_1. According to the mobile user trajectory, i.e., the tuple (old AP, old CH), the new visited AP assigns the new CH to the mobile user. In our example, when the mobile user enters to the subnet managed by AP_5 with the tuple (AP_4, AP_1), it registers to the CH AP_1. The new state of the mobile user becomes (AP_5, AP_1). On the other hand, when the mobile user visits AP_5 while having (AP_6, AP_6) as the current state, it registers to the CH AP_6. The new mobile user state is therefore (AP_5, AP_6). This simple example shows clearly the dynamic and distributed properties of the DCA algorithm. More formally, the distributed clustering algorithm using the INX routing metric is described by the pseudocode in Fig. 4.

Note that the cluster formation in DCA depends on the routing metric used for path selection. For instance, using the hop-count metric, the cluster managed by AP_1 includes AP_9 and AP_{10} in addition to the set provided by the INX-based routing (see Fig. 3). This is determined according to the relation in (9).

It is worth noting that the DCA algorithm needs all the APs to maintain the information regarding the distances between them as well as the number of interferer links of the paths connecting them (i.e., the interference matrix I, as described in Section 3.1). To do so, we can either setup this information statically in advance at each AP when the WMN is deployed or use a distance vector routing protocol such as AODV [34] or DSDV [35] to discover and disseminate it. As APs in WMNs are not mobile components, each one maintains a quasi-static table containing this information. Note that, in the case of using the INX metric, we need also to broadcast small probe packets to measure the wireless link quality in terms of packet loss. Then, using these measurements, an AP can compute the INX metric associated to each link using the Eq. (2). The route metric, which is the sum of the link metrics, will be accumulated since the used routing protocol (i.e., AODV or DSDV) forwards

```
%Location registration procedures
IF (Mobile user enters a new subnet)
    Sends a Reg. Req. containing the CH's IP address
    and the interference cost between the GW and the CH;
    New AP checks the existence of the current CH's IP
    address in its routing table;
    IF (the address exists)
        New AP finds the minimum interference path
        to the CH and to the GW;
        IF (INX(newAP,CH) < INX(newAP,GW))
        & (INX(GW,CH) + INX(CH,newAP) ≤ I_max)
            Perform a CH registration: new AP added to the cluster;
        ELSE
            Perform a GW registration: new cluster formation;
            CH renewal: new AP becomes the new CH;
        ENDIF
    ENDIF
ENDIF
```

Fig. 4. The DCA algorithm using INX as routing metric.

updates in the network. Note also that, because the probes are broadcast, 802.11 MAC protocol, for example, does not acknowledge or retransmit them.

In this case, the changes in the physical topology are more efficiently handled by the DCA strategy, since it reacts much more quickly than the OSC strategy. Typically, following an AP failure or a new AP addition, the DCA algorithm needs only the distance table or the interference matrix I to be updated at each AP. Then, each AP identifies automatically its corresponding virtual cluster when it operates as a CH. This is achieved without any additional computing complexity due to the distributed nature of the DCA algorithm. In contrast, using the OSC approach, the optimal placement of the clusters must be recalculated centrally at the gateway according to the new topology (i.e., connectivity graph). This operation is time consuming and may take from several hours to several days according to the size of the WMN (i.e., the WMN size in terms of number of APs and the number of wireless links). Moreover, this information regarding the clusters placement must be disseminated to each AP, which introduces additional signaling cost compared to the DCA algorithm. In view of this, the DCA algorithm presents several advantages compared to the OSC method from operational flexibility point of view.

In the remainder of this paper, we will present an in depth comparison between DCA and OSC strategies based on their RRU costs, load balancing and time complexities. Moreover, these clustering schemes will be compared to the case where clustering is not considered under various mobility and traffic scenarios and using different routing metrics. In addition, we will evaluate the potential performance gains of INX compared to ETX and hop-count in terms of total network throughput, packet loss and delivery probabilities of used links. Our aim is to study the impact of interference on our proposed clustering algorithms.

6. Performance metrics

In this section, we present the performance metrics used to compare the DCA, OSC and WC methods in terms of RRU cost, registration updates cost, data delivery cost and load balancing. Explicit analytical expressions of these metrics are derived in the OSC and WC cases. Simulation results are instead used in the DCA case.

In our analysis, we consider a general two-dimensional (2-D) random walk mobility model [36,37]. Accordingly, a mobile user connected to AP_i moves to one of the neighboring subnets with equal probability p (i.e., $p = \frac{1}{\delta(AP_i)}$, where $\delta(AP_i)$ is the AP_i connectivity degree). Using these probabilities, we construct the transition probability matrix $P = [p_{ij}]$ between pairs of mesh nodes. This matrix is used as an input parameter to solve the ILP problem described in the previous section. Another necessary input parameter used to solve the ILP problem is the interference matrix I, which is computed as shown in Section 3.1. The probability vector (Π_i), $i = 1, \ldots, N$ that the mobile user is located at the subnet AP_i is also used as an input parameter. This probability vector is derived by resolving the following system since the above-mentioned process $\{Y(t), t \geqslant 0\}$, where $Y(t) = AP_i$ represents the user's location at time t, is Markovian.

$$\Pi P = \Pi \text{ and } \sum_{i=1}^{N} \Pi_i = 1. \tag{11}$$

6.1. Data delivery cost

The data delivery cost is defined as the average interference cost of a path used to forward data packets between the gateway and the mobile user. It is representative of the average number of blocked links inside the WMN when a packet is delivered from the gateway to the current serving AP (i.e., in the downlink direction). Note that in our study, we consider only the packet delivery cost of the downlink traffic. This is because the packet delivery cost in the uplink direction (i.e., from the mobile user to the gateway) is the same for all the approaches (i.e., DCA, OSC and WC) since the path with minimum number of interferer links is always used in this direction if the INX metric is used for route selection. So, we omit this term as it is the same for all the compared approaches.

In both DCA and OSC methods, packets destined to the mobile user have to pass-through the CH due to the clustering process. Hence the data delivery cost metric for OSC can be simply given by substituting (8) in (7):

$$\mathit{Data\_Delivery\_Cost}(\mathrm{OSC}) = \frac{1}{N_l} \times \sum_{i=1}^{N} (\Pi_i \times \sum_{j=1}^{N} b_{j,i} \times [I(GW, j) + I(j, i)]), \tag{12}$$

where $b_{j,i}$ is an output of the ILP problem resolution defined in Section 5.1.

In the case where clustering is not considered, packets are delivered using the minimum cost along the path between the gateway and the mobile user. Hence, the data delivery cost is given by:

$$\mathit{Data\_Delivery\_Cost}(\mathrm{WC}) = \frac{1}{N_l} \times \sum_{i=1}^{N} \Pi_i \times I(GW, i). \tag{13}$$

6.2. Registration updates cost

It denotes the delivery cost of the registration updates messages when a handoff occurs. In other words, it is the average interference cost of a path used to exchange the signaling messages needed to track the user mobility. It is representative of the number of blocked links due to the exchange of signaling messages in the WMN when the mobile user moves to a new AP. In both DCA and OSC methods, a local registration (i.e., *CH registration*) is required as long as the mobile user remains in the same cluster. Otherwise, a *GW registration* is performed.

Considering the OSC method, the registration updates cost can be expressed as follows by replacing (6) in (5):

$$Reg\_Update\_Cost(\mathrm{OSC}) = \frac{1}{N_l} \times \sum_{i=1}^{N} \left(\Pi_i \times \sum_{j=1}^{N} P(i,j) \times \sum_{k=1}^{N} b_{k,j} \times [I(j,k) + (b_{k,j} - b_{k,i}) I(k, GW)] \right). \quad (14)$$

When the clustering policy is not considered, the registration updates cost is given by:

$$Reg\_Update\_Cost(\mathrm{WC}) = \frac{1}{N_l} \times \sum_{i=1}^{N} \Pi_i \times Update\_Cost(i), \quad (15)$$

where $Update\_Cost(i)$ is the cost of registration updates when the mobile user leaves the AP_i. It can be written as follows:

$$Update\_Cost(i) = \sum_{j=1}^{N} P(i,j) \times I(AP_j, GW). \quad (16)$$

Alternatively, the registration updates cost in the WC case can be also calculated as follows:

$$Reg\_Update\_Cost(\mathrm{WC}) = \frac{1}{N_l} \times \sum_{i=1}^{N} \Pi_i \times Update\_Cost'(i),$$

where $Update\_Cost'(i)$ represents rather the cost of registration updates when the mobile user moves to the AP_i, which is simply equal to $I(AP_i, GW)$. Hence, the registration updates cost can be written also as follows:

$$Reg\_Update\_Cost(\mathrm{WC}) = \frac{1}{N_l} \times \sum_{i=1}^{N} \Pi_i \times I(AP_i, GW). \quad (17)$$

Since the interference matrix I is symmetric, it is easy to see that, in the WC case, the registration updates cost shown in (17) equals the data delivery cost shown in (13). This is simply because both packet delivery and registration updates are always performed with the same node, which is the gateway.

6.3. RRU Cost

It is the interference cost of radio resources used inside a WMN where the profile of its mobile users are characterized by the pair (λ, μ). λ measures the average rate of traffic exchanged by each mobile user and μ describes the user mobility in the WMN. The expression of the *RRU_Cost* for the OSC and WC cases can be simply obtained by replacing (12) and (13) in (4), and (14) and (17) in (4), respectively.

Recall that in the particular case where clustering is ignored (i.e., the WC case), the data delivery cost equals the registration updates cost, as demonstrated above. As such, we have:

$$\begin{aligned} RRU\_Cost(\mathrm{WC}) &= Data\_Delivery\_Cost(\mathrm{WC}) \\ &= Reg\_Update\_Cost(\mathrm{WC}) \\ &= \frac{1}{N_l} \times \sum_{i=1}^{N} \Pi_i \times I(AP_i, GW). \end{aligned} \quad (18)$$

6.4. Load balancing cost

The load balancing (LB) is defined as the variance of the traffic load handled by the different clusters. The lower the variance we get, the more efficient the load balancing is. Since the WC method does not use any clustering scheme, we only give here the expression of the load balancing for the OSC method. Using the DCA scheme, this performance metric will be derived through simulations.

Let us denote by w_i the proportion of traffic handled by the cluster identified by its clusterhead AP_i $(i = 1, \ldots, N)$. w_i can be written as follows:

$$w_i^{\mathrm{OSC}} = \sum_{j=1}^{N} b_{i,j} \times \Pi_j \forall i = 1, \ldots, N. \quad (19)$$

Note that according to (19), $w_i \neq 0$ if the AP_i is a CH, otherwise $w_i = 0$. Hence the number K of clusters according to the OSC scheme is:

$$K^{\mathrm{OSC}} = \sum_{i=1}^{N} 1_{w_i}^{\mathrm{OSC}}, \quad (20)$$

where 1_{w_i} is the indicator function of the condition $w_i > 0$. Let the vector $W' = [w'_i]$, $i = 1, \ldots, K$, be the subset of $W = [w_i]$, $i = 1, \ldots, N$, representing only the traffic load of the clusterhead APs. As stated before, the load balancing is defined as the variance of the variable w'_i, $i = 1, \ldots, K$. Note that $\sum_{i=1}^{K} w'_i = 1$. Hence $E[W\prime] = 1/K$ and the load balancing between clusters according to the OSC approach is given by:

$$\begin{aligned} LB(\mathrm{OSC}) &= \frac{1}{K^{\mathrm{OSC}}} \left(\sum_{i=1}^{K^{\mathrm{OSC}}} \left(w_i'^{\mathrm{OSC}} \right)^2 \right) - \left(\frac{1}{K^{\mathrm{OSC}}} \right)^2 \\ &= \frac{1}{K^{\mathrm{OSC}}} \left(\sum_{i=1}^{N} \left(w_i^{\mathrm{OSC}} \right)^2 \right) - \left(\frac{1}{K^{\mathrm{OSC}}} \right)^2. \end{aligned} \quad (21)$$

7. Performance evaluation

We conducted extensive simulations to evaluate the performance of our clustering algorithms (i.e., OSC and DCA) as well as the proposed interference-aware routing metric INX. The section is divided into two parts: INX-based and cluster-based simulations.

7.1. INX-based simulations

In this set of experiments, we studied the performance of our proposed INX metric in terms of total network throughput, packet loss rate and cumulative distributed function (CDF) of delivery probabilities for used links.

ETX and hop-count routing metrics are used as baseline examples for the performance comparison. To achieve this, a simulation model has been developed using ns-2 [30].

The simulated network consists of 40-node arbitrary WMN with an average connectivity degree equals to 7.9. We modified the AODV routing protocol [34] to include the ETX and INX metrics. Note that an AODV-like routing protocol is suitable to use with ETX and INX because these metrics need to collect the value of each link on the path.

Two cases are distinguished. First, at any time, only one connection (flow) is active. Studying this basic case enables us to explore the impact of inter-hop interference on the network throughput. Note that the throughput is measured by the number of data bytes successfully received at the gateway node. Then, we investigated the case of multiple simultaneous connections, which is more likely to be the case in real networks.

7.2. Basic performance

Before we discuss the performance of our metric with multiple simultaneous connections, it is essential to establish a baseline with only one active connection at any time.

In our simulations, we carried out a 3-min TCP connection between the gateway and the remaining nodes. During each TCP connection, the sender transmits as much packets as it could to the gateway (in our simulations, each source transmitted 512-byte data packets at a rate of 11 Mbps). It has been demonstrated in [38] that 1-min TCP connections were of sufficient length to overcome start up effects. Hence, between consecutive TCP connections, we set a 1-min idle period in which the adopted routing metric is calculated and serves to establish the route for the next arriving connection.

Table 1 shows the average network throughput when using the three underlying routing metrics when only a single TCP connection is active at any time. The values are arranged by hop-count. We can observe that ETX provides slightly better results than our proposed INX metric. This is because both INX and ETX take into account the packet loss rate of the wireless links. However, as a distinguishing feature from ETX, INX takes into account the impact of a selected path on the subsequent connections. This is clearly shown in Fig. 5, which compares the cumulative distributed function (CDF) of delivery probabilities for the links used by INX, ETX and hop-count. The median delivery probability is 0.85 for both INX and ETX metrics, meaning that these metrics often use links with loss rate of 15% or more. Whereas, in the hop-count metric, the median delivery probability is 0.78, which means that this metric often uses links with loss rates of 22% or more.

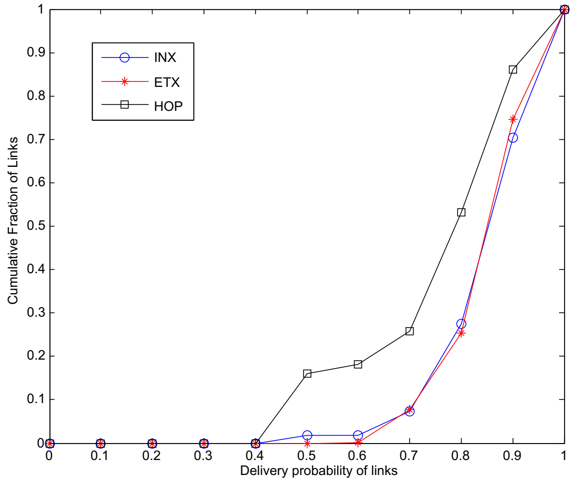


Fig. 5. CDF of delivery probabilities of used links when only one connection is active at any time.

We can see that INX may favor the use of lower quality links to reduce the resulting interference with eventual new arriving connections. Since there is only one active connection in the network at any time, one could expect no improvement of INX over ETX. Instead, we would expect that ETX outperforms INX, since ETX selects always the best current path as opposed to INX. This is indeed confirmed by the results in Table 1, where ETX provides slight improvement (around 3%) in average throughput than INX.

Table 1
Average TCP throughput and packet loss rate between the gateway and each node in the network when only one connection is active at any time

Hops	Number of pairs			Throughput (Kbps)			Packet loss rate (%)		
	INX	ETX	HOP	INX	ETX	HOP	INX	ETX	HOP
1	10	10	14	1921	1921	1509	0.17	0.17	0.58
2	12	12	9	1048	1070	473	0.74	0.49	2.08
3	8	8	12	605	624	261	0.51	0.45	2.9
4	9	9	4	455	469	369	2.12	1.35	4.1

7.3. Simultaneous connections

In the previous scenario, only a single TCP connection was active at any time. This is unlikely to be the case in real networks. In this scenario, we consider multiple simultaneous connections. We compare the network throughput provided by the different routing metric: INX, ETX and hop-count.

As before, each TCP connection lasts 3 min. Instead, the inter-arrival time of connections is varied. In the previous case, the inter-arrival time was 4 min. Consequently, there is at most one active connection in the network at any time. Reducing the inter-arrival time increases the number of simultaneous active connections. For instance, setting the inter-arrival time equals to 1 minute results in the presence of 3 simultaneous active connections at any time.

Fig. 6 plots the throughput of one connection randomly selected among 39 possible ones before and after the arrival of a new second connection, and when using the different routing metrics. We can observe that the hop-count always provides the poorest results since it selects the shortest available path regardless the quality of links in that path. With ETX, the network throughput increases significantly compared to the hop-count metric. On the other hand, our INX metric improves considerably the network throughput compared to ETX in this scenario. Indeed, INX can attribute routes for arriving connections progressively

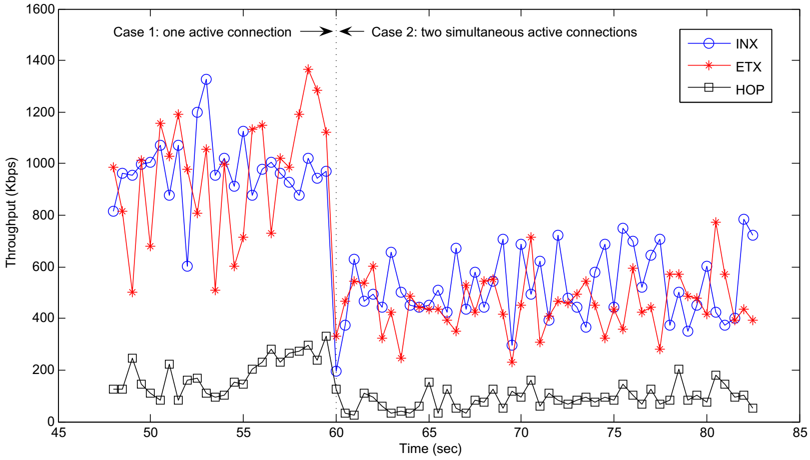


Fig. 6. Throughput of one connection before and after the arrival of a new second connection.

Table 2
Average TCP throughput and packet loss rate between the gateway and each node in the network when two simultaneous active connections are present at any time

Hops	Number of pairs			Throughput (Kbps)			Packet loss rate (%)		
	INX	ETX	HOP	INX	ETX	HOP	INX	ETX	HOP
1	7	7	14	997	838	663	2.4	6.4	18
2	7	5	8	793	496	153	1	2.35	31.5
3	6	7	12	620	525	49	0.8	3.22	56.5
4	8	8	4	559	494	57	1.09	2.45	56.15
5	8	9		408	351		2.85	4.8	
6	2	2		373	319		2.7	4.55	

while preserving good paths for the subsequent flows. This result is important since this situation happens more likely in real networks. Compared to ETX, the INX metric takes more advantage from the parallelism in the network. It may attribute a lower quality path than that selected by ETX for an arriving connection in order to keep useful resources for the subsequent connections. As such, the total network throughput is improved, as shown in Table 2. Indeed, the improvement achieved by INX in average throughput can attain 38% and 92% compared to ETX and hop-count, respectively.

Fig. 7 shows the CDF of delivery probabilities for the links used by INX, ETX, and hop-count when two simultaneous connections are active at any time in the WMN. The median delivery probability is 0.77 for INX, 0.7 for ETX and 0.49 for the hop-count metric. This means that INX, ETX and hop-count often use links with loss rates of 23% or more, 30% or more, and 51% or more, respectively. We can see that with INX, the second arriving connection find better routes than that when using ETX as opposed to the Fig. 5. This is again because INX takes into consideration the inter-connection interference when selecting paths and accordingly preserves good links for subsequent connections.

Fig. 8 shows the average packet loss rate during the session as function of number of hops for both scenarios and when using the different routing metrics. We can see that

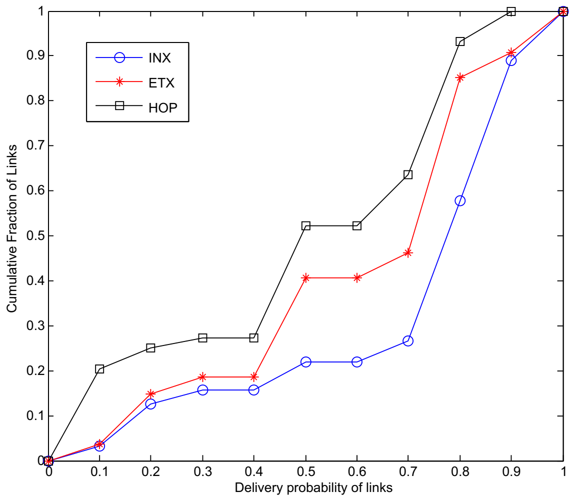


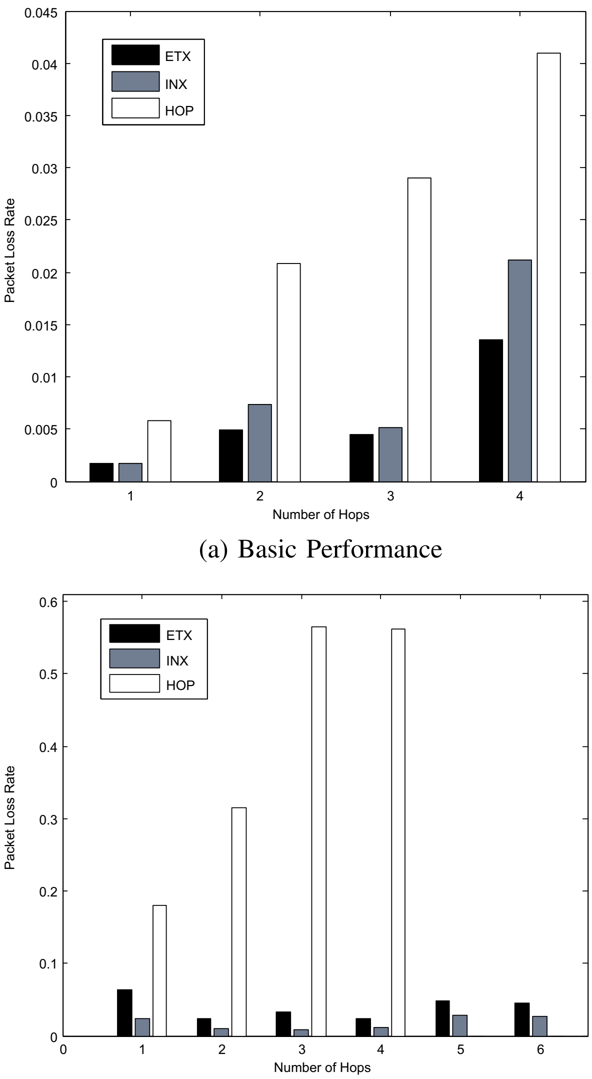
Fig. 7. CDF of delivery probabilities of used links when two simultaneous active connections are present at any time.

both INX and ETX reduce the packet loss during a session compared to the hop-count metric, especially for the case of multiple simultaneous connections. INX allows the lowest packet loss rate since the interference caused by the subsequent connections is taken into account.

7.4. Cluster-based simulations

In the earlier subsection, we evaluated the efficiency of our proposed INX metric against both ETX and hop-count using two different scenarios. In this subsection, we investigate the use of this metric in the clustering process. Our aim is to study the impact of addressing efficiently the interference effect between links on our clustering algorithms (i.e., OSC and DCA), and compare the obtained results with the conventional case, where the hop-count metric is used.

As stated before, since no clustering approaches have been previously proposed to deal with the mobility man-



(a) Basic Performance

(b) Simultaneous Connections

Fig. 8. Average packet loss rate during a session.

agement in WMNs, the OSC and WC cases are used as baselines to which the DCA improvements are compared. The simulations are performed using both ns-2 [30] and a new discrete-event simulation tool developed under Matlab [31].

The analysis is achieved based on arbitrary meshed topologies. We considered different network sizes: $N = 20$, $N = 60$ and $N = 100$ APs, which are representative of small, medium and large single-gateway WMNs. In our simulations, the N APs are distributed randomly in the network area. Then, based on the transmission range R_t and the hearing range R_h, which is taken equal to twice R_t, the connectivity matrix and the interference matrix are derived, respectively. It is worth noting that the resulting connectivity graph is considered only if it is connected, i.e., there is at least a path that connects each AP to the gateway. Otherwise, we keep generating random topologies until the graph connectivity condition is satisfied. Once a valid arbitrary meshed topology is obtained, we can also modify the connectivity graph by increasing R_t. Obviously, the resulting graph is always connected and in doing so, we increase the graph degree $\bar{\delta}$. The rational behind this is to study the impact of the graph degree $\bar{\delta}$ on the evaluated metrics (i.e., RRU cost, data delivery cost, registration updates cost and load balancing).

We recall that in our study, we use a random walk mobility model. Accordingly, a mobile user connected to AP_i moves to one of the neighboring subnets with equal probability p. In our experiments, the sojourn time $1/\mu$ within an AP subnet is set equal to 10, 100 or 1000 seconds to represent very fast, fast or slow mobile users, respectively. In addition, we used different values of λ to represent different mobile user loads (i.e., light, moderate and heavy-traffic mobile user loads). Specifically, the values of λ and the associated quantity of downloaded traffic per user and per day are reported in Table 3. To get an estimate from λ of the downloaded traffic per user, we assume that the average packet size is equal to 460 bytes (see [39]).

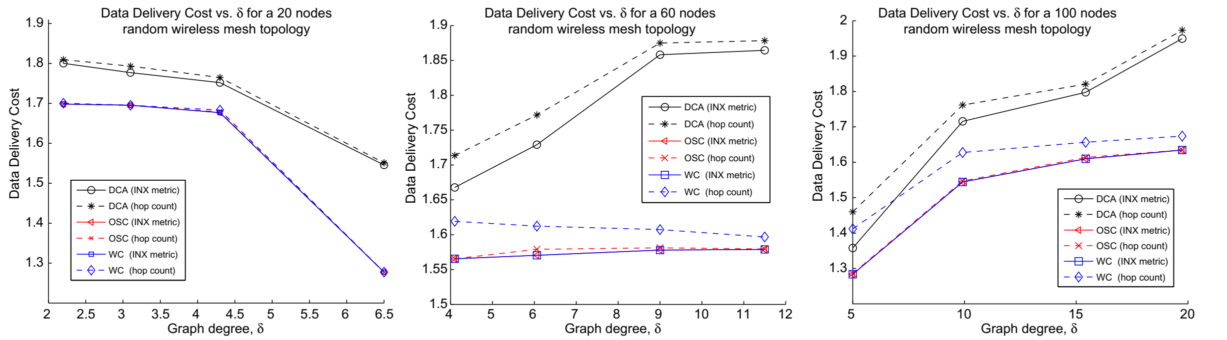
It is worth noting that in our simulations, in order to take into account the interference effect between links, each arbitrary meshed topology is simulated first under ns-2 to compute the packet loss rate of each available wireless link. Based on these results, we calculate the ETX value of each link and then derive the INX-based routing matrix using (2). This matrix is then used as input for our discrete-event simulation tool to derive the above-mentioned performance metrics.

As mentioned before, we will first compare the DCA and OSC clustering strategies based on their RRU costs and load balancing using both INX and the hop-count routing metrics. Subsequently, these clustering schemes are compared under various mobility and traffic scenarios, which are characterized by the pair (λ, μ) to the case where clustering is not considered in order to evaluate the benefit of clustering in WMNs.

The results are reported in Figs. 11–15. Before delving into the exploration of these graphs, let us start by analyzing the results regarding the data delivery and registration updates costs associated with the derived RRU cost according to the different strategies. Recall that the RRU cost of a mobile user is simply a weighted sum of two terms: the data delivery and registration updates costs (see (4)).

Table 3
Quantity of downloaded traffic per user per day

λ (packets/s)	0.001	0.01	0.1	0.5	1	2	10
Traffic (Mbits/day)	0.318	3.18	31.8	159	318	636	3180



(a) $N = 20$ (b) $N = 60$ (c) $N = 100$

Fig. 9. Data delivery cost in an arbitrary wireless mesh network.

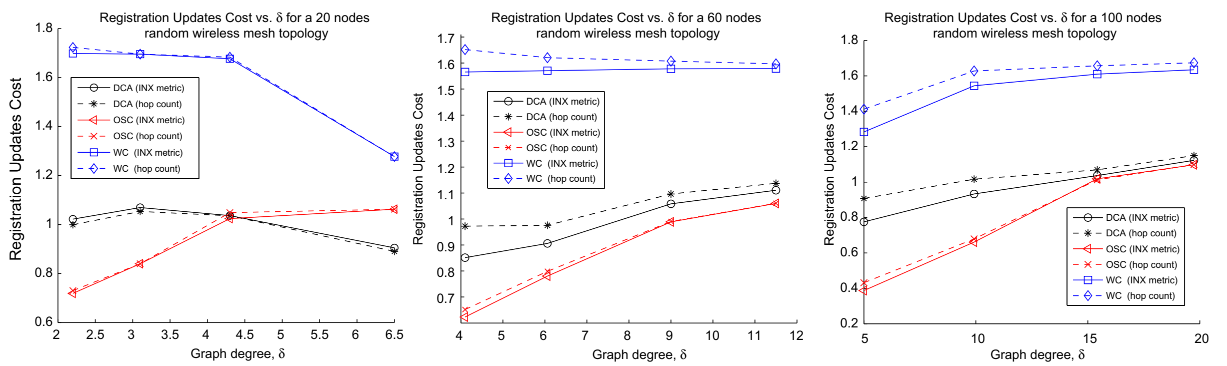
7.5. Data delivery cost results

Fig. 9 presents the simulation results regarding the data delivery cost using both the hop-count and the INX routing metrics. Recall that the data delivery cost is representative of the average number of blocked links due to the transmission of a data packet from the gateway to the current serving AP. The *X*-axis of the graphs represents the average node degree $\bar{\delta}$ in the network (i.e., graph degree).

The first thing to note is that using the INX metric instead of the hop-count metric, we reduce considerably the data delivery cost, notably when the graph degree is low. This is shown more clearly in Fig. 9b and c, and explains why the shortest path often performs poorly in dense WMNs since it leads often to the use of paths with high interference cost.

We can also observe that the minimum data delivery cost is achieved when clustering is ignored and when the routing metric INX is used. This cost increases with DCA due to the additional cost introduced by the clustering process. However, with the OSC approach, the data delivery cost does not increase compared to the optimal WC case although clustering is adopted. This means that with the OSC approach, the CH associated to a given AP_i is always chosen to be an AP belonging to the route with the minimum interferer links between the AP_i and the gateway. Moreover, we can notice that the cost regarding the hop-count based OSC approach remains the same compared to the case where the routing metric INX is used. This is because the objective function to be minimized in the OSC case corresponds for both routing metrics to the RRU cost, which is defined as the total number of interferer links along the paths utilized by signaling and data packets. In this regard, the optimal solution returned by the ILP resolution and that minimizes the RRU cost in the OSC case, minimizes also the data delivery cost term included therein when using either the INX or the hop-count routing metrics.

Finally, with regard to the impact of the graph degree $\bar{\delta}$, we can see that the data delivery cost decreases with the increase of $\bar{\delta}$ in small WMNs as shown in Fig. 9a. This is because in such networks, shortest paths often offer the minimum interference cost. This cost decreases as it is proportional to the average number of hops between nodes, which decreases with the increase of $\bar{\delta}$. In contrast, the data delivery cost in medium and large WMNs increases with the graph degree $\bar{\delta}$ for both metrics as depicted in Fig. 9b and c, respectively. This is because the number of interferer links increases with $\bar{\delta}$ as the interference range of each node increases.



(a) $N = 20$ (b) $N = 60$ (c) $N = 100$

Fig. 10. Registration updates cost in an arbitrary wireless mesh network.

7.6. Registration updates cost results

Fig. 10 plots the obtained registration updates costs following the RRU minimization according to the different strategies as a function of the graph degree $\bar{\delta}$. As a first observation, we can see that the clustering technique reduces considerably the registration updates cost using either the hop-count or INX metrics, since some expensive *GW registrations* are replaced by low-cost *CH registrations*. Recall that both the OSC and DCA approaches divide the WMN into virtual clusters in order to limit the mobile user registration updates within this local area, instead of communicating with the far away gateway. In this regard, the registration updates cost decreases significantly with OSC and DCA approaches.

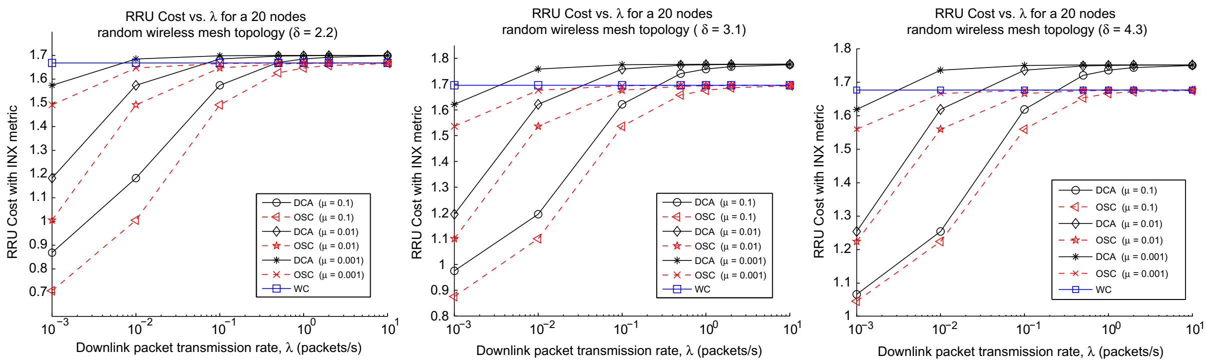
As a second main observation, we can see in Fig. 10a that the INX and hop-count metrics results almost coincide. The reason is that, in a small WMN where the paths are not too long, the shortest path often offers the minimum interference cost.

The third main observation concerns the WMNs with medium and large sizes. Indeed, we can see in Fig. 10b and c that OSC performs the best as it has the lowest registration updates cost. DCA achieves near-optimal performance especially for high values of $\bar{\delta}$. In addition, we can notice from these figures that the registration updates cost of all approaches increases with the graph degree $\bar{\delta}$. This is because the number of blocked links due to interference increases with the increase of $\bar{\delta}$. Finally, we can notice that the performance gap between the hop-based simulation results and the INX-based ones decreases with the increase of $\bar{\delta}$. The reason is related to the fact that short paths, which are created when $\bar{\delta}$ is high do not cause heavy interference.

7.7. RRU cost results

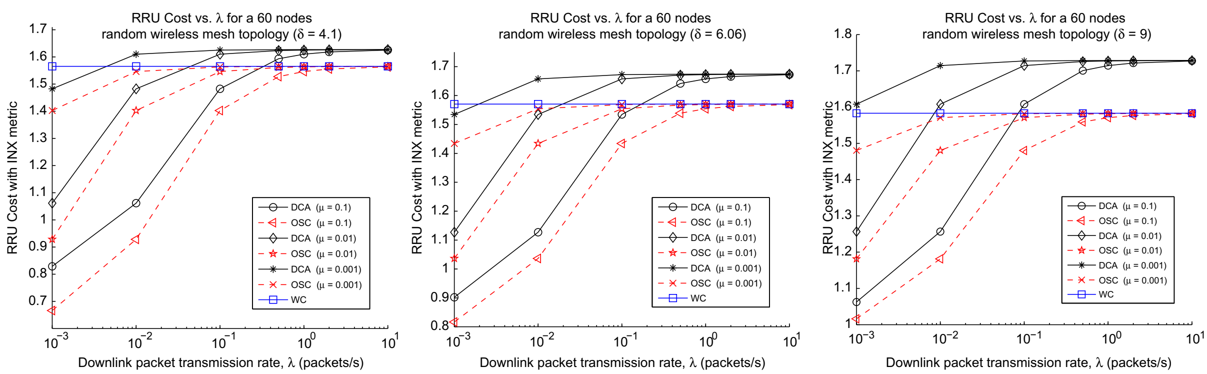
In this subsection, we analyze the results regarding the RRU cost provided by the different strategies under various scenarios. To achieve this, we will proceed as follows. We will first identify the cases where clustering is useful. Then, when clustering is helpful, we will compare between the OSC and DCA clustering approaches using INX as well as the hop-count routing metric.

First, we recall that the RRU cost is simply a weighted sum of the already discussed data delivery and registration



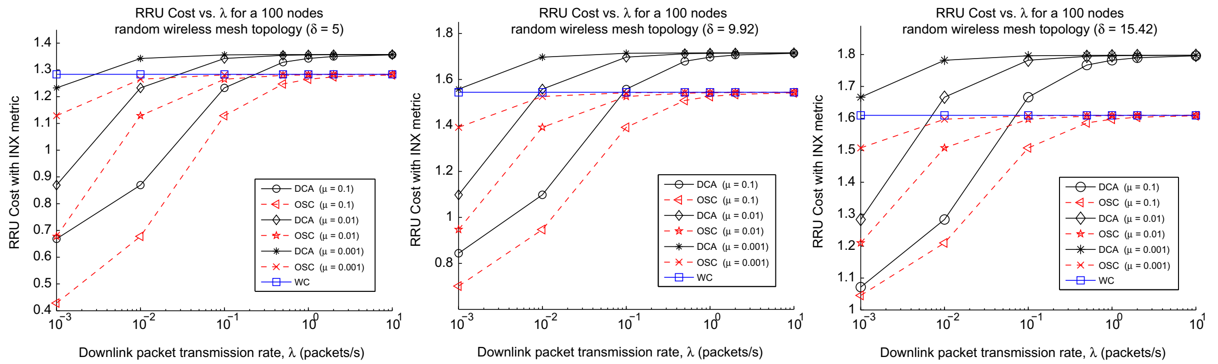
(a) Graph degree $\bar{\delta} = 2.2$ (b) Graph degree $\bar{\delta} = 3.1$ (c) Graph degree $\bar{\delta} = 4.3$

Fig. 11. Radio resource utilization (RRU) cost in a 20-node arbitrary wireless mesh network with INX metric.



(a) Graph degree $\bar{\delta} = 4.1$ (b) Graph degree $\bar{\delta} = 6.0\ 6$ (c) Graph degree $\bar{\delta} = 9$

Fig. 12. Radio resource utilization (RRU) cost in a 60-node arbitrary wireless mesh network with INX metric.



(a) Graph degree $\bar{\delta} = 5$ (b) Graph degree $\bar{\delta} = 9.9\ 2$ (c) Graph degree $\bar{\delta} = 1\ 54\ 2$

Fig. 13. Radio resource utilization (RRU) cost in a 100-node arbitrary wireless mesh network with INX metric.

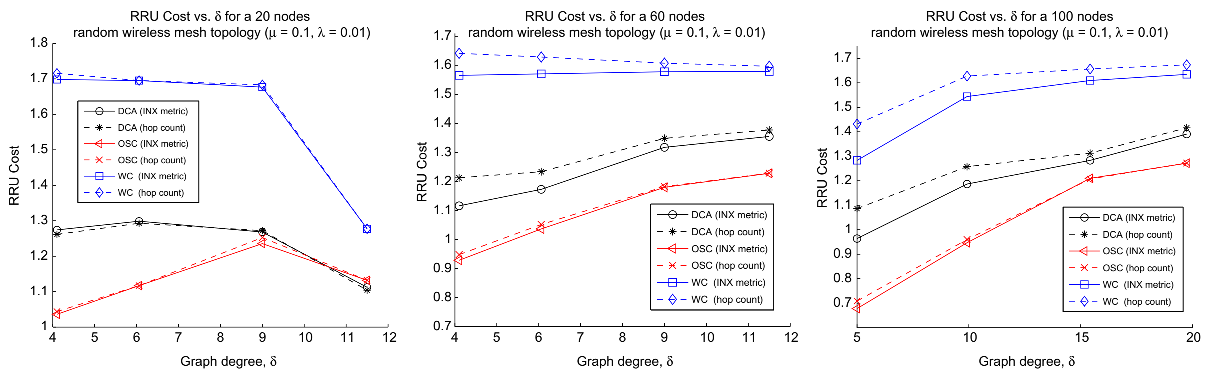
updates costs. In the WC case, the RRU cost equals the data delivery cost and the registration updates cost. Hence, as shown in Figs. 11–13, the RRU cost with the WC approach is insensitive to the mobile user profiles (i.e., λ and μ), and is a constant metric that depends only of the physical topology (i.e., number of nodes and graph connectivity properties) as it will be illustrated in Fig. 14.

In the previous subsections, we have proven that ignoring clustering enables the lowest data delivery cost. Moreover, we have shown that the main aim of clustering is to reduce the registration updates cost, probably at the expense of a slight increase of the data delivery cost for the DCA scheme. Intuitively, when the amount of signaling messages that a mobile user generates due to its mobility is negligible, i.e., the registration updates cost has a minor impact on the total RRU cost, then clustering is not helpful. In other words, if $2\mu \times m_{\mathrm{sig}} \ll \lambda \times m_{\mathrm{data}}$, clustering does not provide any gain. In view of this and as a first main finding, we can see through Figs. 11–13 that when λ and $1/\mu$ are high, the WC approach stands out as the best solution. Clustering becomes interesting only for small and moderate values of λ and $1/\mu$. This range of μ implies short sojourn times of the mobile users in each AP subnet, which results in frequent hand offs and thus increasing the signaling traffic load. Accordingly, the registration updates cost is no longer a negligible cost and thus both OSC and DCA achieve a significant gain compared to the WC case. This is likely to be the case in cities, where users are often mobile.

It is worth noting that in the OSC and DCA cases, the RRU cost increases progressively with the increase of λ and $1/\mu$ from the registration updates cost to attain the data delivery cost. This explains why the RRU cost in the DCA case exceeds that of WC when λ and $1/\mu$ have large values. In fact, in this range of λ and μ, the RRU cost is equivalent to the data delivery cost as it is a dominant cost. Moreover, we recall as shown in Fig. 9 that the DCA data delivery cost exceeds that of the WC case since the packets are not forwarded directly to their destinations and have to transit by their associated CHs through the paths that offer the minimum interference cost between the gateway and the CH as well as between the CH and the visited AP.

In turn, the OSC approach achieves always better results than the WC case. In this context, we point out that the WC strategy can be seen as a particular case of the OSC one where the number of clusters is set equal to 1 and where the gateway is the only CH. In this regard, both strategies exhibit similar results when λ and $1/\mu$ have relatively large values. This is shown clearly in Figs. 11–13 by the asymptotic behavior of the OSC curves with respect to the WC constant (i.e., horizontal) one.

Let us now focus on the comparison between the OSC and DCA clustering techniques using first the routing metric INX. Beforehand, let us recall that the OSC approach



(a) $N = 20$ (b) $N = 60$ (c) $N = 100$

Fig. 14. Radio resource utilization (RRU) cost comparison in *N*-node arbitrary wireless mesh network with INX and hop-count routing metrics.

provides the optimal RRU cost when the WMN is divided into disjoint clusters. As a distinguishing feature of the DCA approach, clusters are not disjoint, in the sense that the same AP may belong to different clusters. In this case, an AP may be associated to different CHs during the mobile user journey according to its trajectory (see our example in Fig. 2b). In what follows, we will see if the DCA approach gives also reasonable results with respect to the optimal OSC one.

We can observe from Figs. 11–13 that the OSC approach always performs the best by offering a minimal RRU cost. This result is expected since the returned data delivery overhead and registration updates overhead are often minimal with the OSC approach as shown in Figs. 9 and 10, respectively. The tradeoff between these two overheads is achieved through the objective function of our ILP formulation.

By comparing to the results of OSC, we can see that the DCA method could achieve a near-optimal performance especially for high graph degree $\bar{\delta}$ and high or moderate value of users' mobility, which is almost the case in cities (see Figs. 11c, 12c and 13c). However, for low mobility rate, such as in-campus environment, the DCA algorithm becomes more expensive than the WC scheme, as pointed out earlier. The tradeoff between registration updates cost and data delivery cost is in fact achieved in the DCA approach through the cost comparison constraints described in (9) and (10).

In view of this, we can state that the optimal disjoint-clustering method can serve as a lower bound for performance evaluation. However, the resolution of the associated ILP problem is time consuming and may not react efficiently to the physical network topology changes (AP failure, new AP addition, etc.), as opposed to the near-optimal DCA approach, which adapts instantaneously to network changes.

To get an insight into the computation time needed by the clustering methods (i.e., OSC and DCA) in order to identify the set of clusters and their associated CHs, let us consider Table 4. The measurements are performed on a PC with 3.2 Ghz of CPU and 2.00 GB of RAM for a single-gateway arbitrary WMN with N nodes varying between 7 and 91. The reported results show that the DCA algorithm achieves great time saving compared to the OSC approach notably when N is high. The time needed to resolve the OSC optimization clustering problem increases dramatically with the network size N since the number of variables and equations in the ILP formulation increases exponentially with N. In contrast, with DCA, the clusters are set up automatically and in a distributed manner according to both conditions regarding the data delivery cost and the registration updates cost. In this case, the changes in the physical topology are more efficiently handled by the DCA strategy, since it reacts much more quickly than the OSC strategy. Typically, following an AP failure or a new AP addition, the OSC-based network must be reconfigured according to the new optimal cluster placement returned by the ILP resolution. This operation is time consuming and may last several hours up to several days if the network size is large as opposed to the instantaneous reaction in the DCA case.

Table 4
Computation times of DCA and OSC methods

N	7	19	37	61	91
DCA Time (seconds)	0.31	0.98	1.97	7.81	26.05
OSC Time (seconds)	0.047	0.187	33.218	674.078	6570.77

Now, we turn to the case when both the hop-count and INX metrics are used for RRU cost comparison. Fig. 14 shows the RRU cost as function of the graph degree $\bar{\delta}$ for small, medium and large WMNs.

As a first observation, we can see that either OSC or DCA approaches with the hop-count routing metric achieves significant gains compared to the traditional case where clustering is not used (i.e., the WC scheme). We can also observe that, by using the INX metric, more gains can be achieved since the interference effect between links is taken into account during the clustering process.

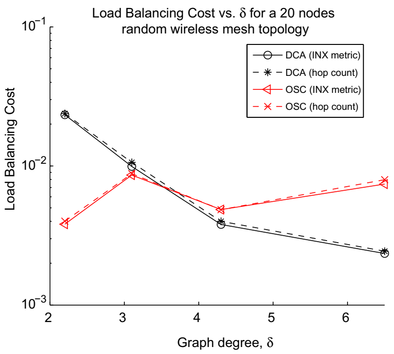
In addition, we can see that the DCA approach provides near-optimal results compared to the optimal ones given by the OSC method, especially for high graph degree. Moreover, from Fig. 14b and c, we can observe that the performance gap between the hop-based routing and the INX-based one decreases with the increase of $\bar{\delta}$. This is because short paths, which are created when $\bar{\delta}$ is high, do not cause heavy interference.

In view of this, two main finding can be highlighted in our paper. First, clustering with the hop-count routing metric is useful and allows a great saving in terms of the aforesaid performance metrics (i.e., registration updates cost, data delivery cost and RRU cost). Second, we showed that by taking into account the interference effect between links, we can further improve the performance of our clustering algorithms and increase the gain initially observed with the hop-count metric. As a result, more profits to the service provider from the existing infrastructure can be realized.

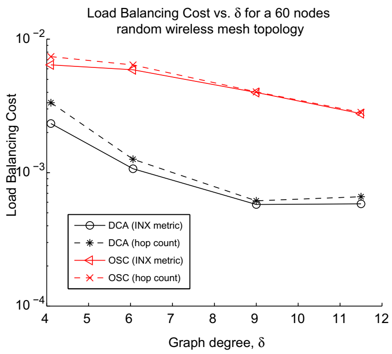
7.8. Load balancing results

Finally, we conclude this paper by studying the load balancing in the WMN following the utilization of the OSC and DCA clustering schemes. Fig. 15a–c illustrate the load balancing for arbitrary meshed topologies when $N = 20$, $N = 60$ and $N = 100$, respectively. Recall that this metric is defined as the variance of the traffic load handled by the different clusters. We can observe from these figures that the DCA algorithm often provides the lowest load balancing cost using either the hop-count or INX routing metric. This is simply because the number of clusters in the DCA approach is much greater than in the OSC case. In addition, the clusters in DCA are formed in distributed manner compared to the centralized way in the OSC case. Typically, each AP can act as a CH in the DCA case. As such, the traffic is better distributed and balanced among the network APs. As a result, we can deduce that the traffic load in DCA is distributed more efficiently among the clusters than in the OSC case reducing thus the probability of congestions in WMNs.

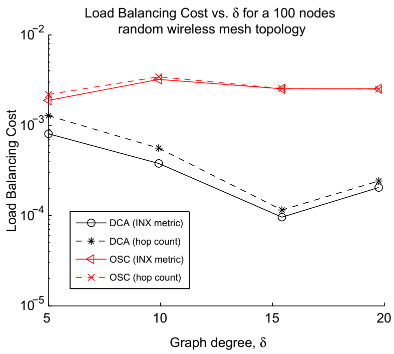
Now, we turn to the case where the hop-count routing is used. Two main observations can be highlighted. First,



(a) $N = 20$



(b) $N = 60$



(c) $N = 100$

Fig. 15. Load balancing in *N*-node arbitrary wireless mesh network with INX and hop-count routing metrics.

considering the OSC approach, we can observe that the hop-based load balancing almost coincide with the INX-based results for high graph degree. This is because the optimal cluster configuration that minimizes the RRU cost in such highly-connected networks is almost the same using either the hop-count or the routing metric INX. This result confirms the simulations plotted in Figs. 9, 10 and 14 when $\bar{\delta}$ is high.

The second main observation concerns the DCA approach. Indeed, we can see that the hop-based simulation results perform worse compared to the INX ones in both medium and large WMNs. This is related to the fact that clusters using the hop-count metric are often large than that obtained when using INX (see Fig. 3 for example). As such, the number of clusters in the INX-based DCA is greater than in the hop-based case. Consequently, the traffic distribution among the network APs can be enhanced when we take into consideration the interference effect among neighboring links. This result is important since the load balancing in WMN remains a critical issue.

8. Conclusion

In this paper, we investigated the radio resource utilization (RRU) efficiency in wireless mesh networks while considering both the mobility of users and the interference effect among neighboring links. We first propose a new interference-aware routing metric, called INX, which aims at improving the total network throughput by providing more opportunity to subsequent arriving connections to be satisfied. Second, we proposed two clustering schemes (i.e., the distributed clustering algorithm, DCA, and the optimal static clustering, OSC) to improve the resource utilization in such networks. Based on extensive simulations and an ILP problem formulation, we proved that our proposed clustering schemes achieve significant gains in terms of RRU cost and load balancing. Specifically, the gain introduced by the clustering schemes depends on the routing policy. In our analysis, we showed that additional gain can be achieved using the proposed INX metric instead of the conventional hop-count metric. In addition, we showed that the DCA clustering technique handles better the changes in the physical topology subsequent to AP failures or addition. Moreover, we showed that the DCA approach achieves a near-optimal solution in terms of RRU cost compared to the optimal OSC results especially for both large and highly-connected wireless mesh networks. Finally, we showed that the load balancing is well improved in DCA compared to the OSC case particularly when the INX metric is used for route selection.

References

[1] MIT roofnet. <http://www.pdos.lcs.mit.edu/roofnet/>.
[2] Seattle wireless. <http://www.seattlewireless.net/>.
[3] R. Karrer, A. Sabharwal, E. Knighly, Enabling large-scale wireless broadband: The case for TAPs, in: Proc. HotNets, November 2003.
[4] I.F. Akyildiz, X. Wang, W. Wang, Wireless mesh networks: a survey, Comput. Netw. 47 (4) (2005) 445–487.
[5] D. De Couto, D. Aguayo, J. Bicket, R. Morris, A high throughput path metric for multi-hop wireless routing, in: Proc. ACM MobiCom'03, San Diego, CA, USA, September 2003.
[6] A.B. McDonald, T.F. Znati, A mobility-based framework for adaptive clustering in wireless ad-hoc networks, IEEE J. Select. Areas Commun. 17 (8) (1999) 1466–1487.
[7] C. Arboleda, N. Nasser, Cluster-based routing protocol for mobile sensor networks, vol. 191, in: Proc. IEEE QShine'06.
[8] J.H. Li, M. yu, R. Levy, A. Teittinen, A mobility-resistant efficient clustering approach for ad hoc and sensor networks, SIGMOBILE Mob. Comput. Commun. Rev. 10 (2) (2006).
[9] D.J. Baker, A. Ephremides, The architectural organization of a mobile radio network via a distributed algorithm, IEEE Trans. Commun. 29 (11) (1981) 1694–1701.
[10] A. Ephremides, J.E. Wieselthier, D.J Baker, A design concept for reliable mobile radio networks with frequency hopping signaling, in: Proc. IEEE, vol. 75(1) (1987), pp. 56–73.
[11] A.K. Parekh, Selecting routers in ad-hoc wireless networks, in: Proc. ITS, 1994.
[12] C.R. Lin, M. Gerla, Adaptive clustering for mobile wireless networks, IEEE J. Select. Areas Commun. 15 (7) (1997) 1265–1275.
[13] N.H. Vaidya, P. Krishna, M. Chatterjee, D.K. Pradhan, A cluster-based approach for routing in dynamic networks, ACM Comput. Commun. Rev. 27 (2) (1997).
[14] S. Basagni, Distributed and mobility-adaptive clustering for multimedia support in multi-hop wireless networks, in: Proc. IEEE VTC, 1999.
[15] S. Basagni, Distributed clustering for ad-hoc networks, in: Proc. ACM ISPAN, June 1999, pp. 310–315.
[16] A. Antis, R. Prakash, T. Vuong, D. Huynh, Max–min *d*-cluster formation in wireless ad-hoc networks, in: Proc. IEEE INFOCOM, 2000.
[17] S. Banerjee, S. Khuller, A clustering scheme for hierarchical control in multi-hop wireless networks, in: Proc. IEEE INFOCOM, 2001.
[18] V. Bharghavan, B. Das, Routing in ad-hoc networks using minimum connected dominating sets, in: Proc. IEEE ICC, June 1997.
[19] J. Wu, H. Li, On calculating connected dominating set for efficient routing in ad-hoc wireless networks, in: Proc. Workshop on discrete

algorithms and methods for mobile computing and communications, 1999.
[20] Y. Chen, A. Liestman, Approximating minimum size weakly-connected dominating sets for clustering mobile ad-hoc networks, in: Proc. ACM MobiHoc, June 2002.
[21] J. Wong, R. Jafari, M. Potkonjak, Gateway placement for latency and energy efficient data aggregation, in: Proc. IEEE LCN, 2004.
[22] R. Chandra, L. Qiu, K. Jain, M. Mahdian, Optimizing the placement of integration points in multi-hop wireless networks, in: Proc. IEEE ICNP, 2004.
[23] Y. Bejerano, Efficient integration of multi-hop wireless and wired networks with QoS constraints, IEEE/ACM Trans. Network. 12 (6) (2004) 1064–1078.
[24] B. Aoun, R. Boutaba, Y. Iraqi, G. Kenward, Gateway placement optimization in wireless mesh networks with QoS constraints, IEEE J. Select. Areas Commun. 24 (11) (2006) 2127–2136.
[25] J. Bicket, D. Aguayo, S. Biswas, R. Morris, Architecture and Evaluation of an Unplanned 802.11b Mesh Network, in: Proc. ACM MobiCom'05, Cologne, Germany, September 2005.
[26] J. Bicket, Bit-rate selection in wireless networks, Master's thesis, Massachusetts Institute of Technology, February 2005.
[27] D. De Couto, D. Aguayo, J. Bicket, R. Morris, A high throughput path metric for multi-hop wireless routing, in: Proc. ACM MobiCom'03, San Diego, CA, USA, September 2003.
[28] K. Xu, M. Gerla, S. Bae, Effectiveness of RTS/CTS handshake in IEEE 802.11 based ad hoc networks, Ad Hoc Network J. 1 (1) (2003).
[29] K. Medepalli, F.A. Tobagi, Towards performance modeling of IEEE 802.11 based wireless networks: a unified framework and its applications, in: Proc. IEEE INFOCOM, Barcelona, Spain, April 2006.
[30] ns2 home page, <http://www.isi.edu/nsnam/ns>.
[31] Matlab home page, <http://www.mathworks.com>.
[32] C.E. Koksal, H. Balakrishnan, Quality-aware routing metrics for time-varying wireless mesh networks, IEEE J. Select. Areas Commun. 24 (11) (2006) 1984–1994.
[33] R. Draves, J. Padhye, B. Zill, Routing in multi-radio, multi-hop wireless mesh networks, in Proc. ACM MobiCom'04, Philadelphia, PA, USA, September 2004.
[34] C.E. Perkins, E.B. Royer, Ad-hoc on-demand distance vector routing, in: Proc. IEEE Workshop on mobile computing and systems and applications, 1999.
[35] C.E. Perkins, P. Bhagwat, Highly dynamic destination sequenced distance-vector routing (DSDV) for mobile computers, in: Proc. ACM SIGCOMM, 1994, pp. 234–244.
[36] K-H. Chiang, N. Shenoy, A 2-D random-walk mobility model for location-management studies in wireless networks, IEEE Trans. Vehicular Technol. 53 (2) (2004) 413–424.
[37] D.N. Alparslan, K. Sohraby, Two-dimensional modeling and analysis of generalized random mobility models for wireless ad hoc networks, IEEE/ACM Trans. Network. 15 (3) (2007) 616–629.
[38] R. Draves, J. Padhye, B. Zill, Comparison of routing metrics for static multi-hop wireless networks, in: Proc. ACM SIGCOMM'04, Portland, OR, USA, September 2004.
[39] Cooperative association for internet data analysis, <http://www.caida.org/>.

Rami Langar received the B.S. degree in telecommunications engineering from the Ecole Superieure des Communications (Sup'-Com), Tunis, Tunisia, in 2001. He received the M.S. degree from the University of Paris 6, Paris, France, in 2002, and the Ph.D. degree from the Ecole Nationale Superieure des Telecommunications de Paris, France, in 2006, both in Network and Computer Science. In 2007 and 2008, he was with the School of Computer Science at the University of Waterloo, Ontario, Canada as a Postdoctoral research fellow. He is currently an Associate Professor of Computer Science at the University of Paris 6. His research interests include 4G visions, mobility management, resource management, wireless mesh networks, vehicular ad-hoc networks, performance evaluation and quality of service support.

Nizar Bouabdallah received the B.S. degree in telecommunications engineering from Ecole Superieur des Communications (Sup'Com), Tunis, Tunisia, in 2001, and the M.S. and Ph.D. degrees in network and computer science from the University of Paris 6, Paris, France, in 2002 and 2004, respectively. He joined Alcatel Research Laboratories, Marcoussis, France, in 2002, while working on his Ph.D. degree. In 2005, he was with the North Carolina State University, Raleigh, NC, USA, as a Postdoctoral Fellow. He is currently a researcher at INRIA (Institut National de Recherche en Informatique et en Automatique). Since February 2007, he has been a Visitor Researcher at the School of Computer Science, University of Waterloo, Canada. His research interests include optical networking, wireless and sensor networks, performance evaluation, network planning and modeling, as well as control and management architectures.

Raouf Boutaba received the MSc. and Ph.D. Degrees in Computer Science from the University of Paris 6, Paris, in 1990 and 1994, respectively. He is currently a Professor of Computer Science at the University of Waterloo. His research interests include network, resource and service management in wired and wireless networks. He is the founder and Editor-in-Chief of the IEEE Transactions on Network and Service Management and on the editorial boards of several other journals. He is currently a distinguished lecturer of the IEEE Communications Society, the chairman of the IEEE Technical Committee on Information Infrastructure and the Director of the ComSoc Related Societies Board. He has received several best paper awards and other recognitions such as the Premier's research excellence award.