Accepted version of:
P. Cappanera, L. Lenzini, A. Lori, G. Stea, G. Vaglini, "Optimal joint routing and link scheduling for real-time traffic in TDMA Wireless Mesh Networks", Elsevier Computer Networks, 57(11), 5 August 2013, pp. 2301–2312, DOI 10.1016/j.comnet.2012.11.021

Optimal joint routing and link scheduling for real-time traffic in TDMA Wireless Mesh Networks

Paola Cappanera
Dipartimento di Sistemi e Informatica
University of Florence, Italy
paola.cappanera@unifi.it

Luciano Lenzini, Alessandro Lori, Giovanni Stea, Gigliola Vaglini
Dipartimento di Ingegneria dell'Informazione
University of Pisa, Italy
{l.lenzini, a.lori, g.stea, g.vaglini}@iet.unipi.it

***Abstract*—We investigate the problem of joint routing and link scheduling in Time-Division Multiple Access (TDMA) Wireless Mesh Networks (WMNs) carrying real-time traffic. We propose a framework that always computes a feasible solution (i.e. a set of paths and link activations) if there exists one, by optimally solving a mixed integer-non linear problem. Such solution can be computed in minutes or tens thereof for e.g. grids of up to 4x4 nodes. We also propose heuristics based on Lagrangian decomposition to compute suboptimal solutions considerably faster and/or for larger WMNs, up to about 50 nodes. We show that the heuristic solutions are near-optimal, and we exploit them to gain insight on the schedulability in WMN, i.e. to investigate the optimal placement of one or more gateways from a delay bound perspective, and to investigate how the schedulability is affected by the transmission range.**

Keywords—Link Scheduling; Routing; Wireless Mesh Networks; Real-time Traffic; Worst-Case Delay

I. INTRODUCTION

Wireless Mesh Networks (WMNs) [1] are a cost-effective technology for providing broadband access at the edge of wireline networks, or in remote, rural, or difficult-to-wire areas. Interference among wireless links with overlapping coverage can be sorted out in either the frequency or the time domain. In the first case, different *channels* are assigned to interfering links, a problem known as *channel assignment.* In the second case, which is the one dealt with in this paper, the full frequency spectrum is given to each link, but interfering links are activated on a Time Division Multiple Access (TDMA [2]) basis. In this case, time is slotted and synchronized, and a *link scheduling* algorithm activates only sets of non-interfering links in the same time slot. Link scheduling algorithms are generally more effective if they take into account the (known or estimated) traffic demand and link scheduling is considered jointly with routing, WMNs generally requiring multi-hop communications. Cross-layer approaches where link scheduling or channel assignment and routing are jointly addressed have been extensively studied [3]-[7] in the past few years.

In the recent past, a growing number of works have envisaged using WMNs for transmitting real-time traffic, e.g. road traffic information [34] , video surveillance [35], etc. Real-time traffic requires the worst-case end-to-end delay (henceforth *WCD* for short) to be below a pre-specified bound or *deadline*. However, comparatively few works so far have taken into account the problem of computing deadline-constrained link *schedules* either given a pre-specified routing plan or jointly with routing. Some (e.g., [24]-[26]) tackle the problem of minimizing the *TDMA delay*, i.e. the sum of the waiting times experienced by a bit that is at the front of its queue at each hop, due to TDMA scheduling. This is, however, only a part (and not necessarily the most relevant one) of the end-to-end delay. Some works ([15]-[18]) aim at guaranteeing a minimum *rate*. This guarantees that the WCD is *finite*, but it does not imply that it is within a pre-specified deadline. Others, finally, aim at optimizing the throughput [19]-[22], or reducing the average delay [33]. While all the above goals are indeed important and worth pursuing, they are not enough to guarantee that pre-specified deadlines are enforced if it is actually possible to do so. For instance, minimizing the TDMA delay, as done in [26], yields schedules that largely violate pre-specified deadlines, even though it is possible to find alternative schedules that do meet them. Our previous works [8]-[10] are actually the first to consider deadline-constrained link scheduling in WMNs, also evaluating different architectural options for flow aggregation. However, routing is left outside the scope of these works, by assuming either a tree network topology, with a single possible path from each node to the network gateway ([8], [10]), or an arbitrary but *given* routing plan, upon which a delay-feasible link schedule is computed ([9]). Tackling the problem of routing and link scheduling separately (e.g., in a cascading approach) leads to a loss in effectiveness. In fact, routing decides *which links* a flow traverses, and link scheduling determines the *capacity* of each link. Capacity-unaware routing may thus select routes that will be congested, and link scheduling on fixed

Accepted version of:
P. Cappanera, L. Lenzini, A. Lori, G. Stea, G. Vaglini, "Optimal joint routing and link scheduling for real-time traffic in TDMA Wireless Mesh Networks", Elsevier Computer Networks, 57(11), 5 August 2013, pp. 2301–2312, DOI 10.1016/j.comnet.2012.11.021

routes cannot explore alternative paths. As a consequence, sets of flows may unnecessarily be declared unschedulable.

In this paper we investigate the problem of *joint* routing and link scheduling, in a cross-layer approach, of leaky-bucket constrained flows that request deadline guarantees. We formulate it as an optimization problem, the *Delay-Aware Routing and Scheduling (DARS) problem*, with the objective of minimizing the maximum deadline violation. When a solution with a negative objective is computed, each flow will follow a route that makes it meet its deadline despite interference. We show that the problem can be optimally solved for networks of up to few nodes (e.g., a 4×4 grid). To allow for larger scales, we propose two suboptimal heuristics, that rely on extrapolating the *link conflict serialization (LCS)* from the DARS. In the LCS, sequences of conflicting link activations are statically precomputed using a coloring approach [32], so as to minimize the longest sequence. In the remaining reduced DARS, the activation of each link is computed jointly with routing, so as to minimize the maximum deadline violation. Once conflicting links are serialized, the reduced DARS problem can be solved optimally for a larger scale (e.g., a 5×5 grid); beyond that scale, optimality has to be traded off for computation time. For this reason, we propose a faster scheme based on a Lagrangian decomposition of the reduced DARS. We show that this heuristic scheme is considerably faster (which allows larger-scale WMNs to be analyzed) and performs close to the optimum. Furthermore, this model can be used to extract useful information related to a WMN, e.g. where to place an Internet gateway node, and whether and when it is profitable to have more than one such node, or again how the schedulability of a set of flows is affected by the transmission range.

The rest of the paper is organized as follows: Section II reports the system model and the problem formulation. In Section III we discuss the properties of the optimal solution and present heuristics. We report performance evaluation results in Section IV, and discuss the related work in more detail in Section V. Section VI concludes the paper.

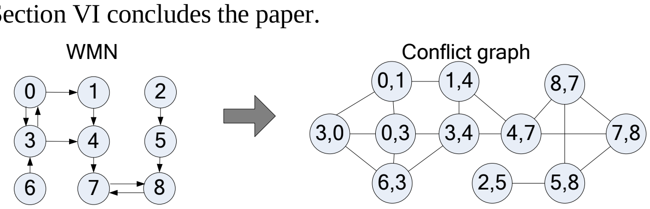

Figure 1. Logical connectivity graph (left) and conflict graph (right) of a WMN.

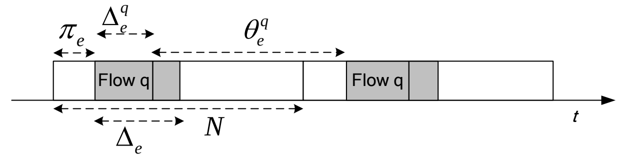

Figure 2. Relevant quantities in link scheduling.

II. System Model and Problem Formulation

The framework developed in this paper relies on basic Network Calculus concepts, i.e. *arrival curve*, *service curve* and *delay bound*. Interested readers can find background in [11], from which we also borrow notation.

We assume that each mesh router is equipped with a *single* time-slotted channel. *Transmission slots* of a fixed duration T_s are grouped into a *frame* of N slots, which is periodically repeated every $N\cdot T_S$ time units. For instance, in 802.16 networks the frame length is usually set to 5ms. Each slot is assigned to a set of non-interfering links through *conflict-free link scheduling*. At every slot, a subset of links may be activated for transmission only if no conflicts occur at the intended receivers. The WMN is modeled through a *connectivity graph*, $G=(V,E)$, whose nodes $V=\{v_1,\cdots,v_n\}$ are mesh routers and whose edges $E=\{e_1,\cdots,e_m\}$ are *directed* links connecting a transmitter to the nodes within transmission range from it. We assume that each link e has a constant transmission rate W_e. The connectivity graph is a *logical* representation of the WMN, which can be derived from the physical WMN topology once the transmit powers, antenna gains, node distances and path loss are known. For instance, in Figure 1 pictures a situation where the transmission range of node 6 is such that 7 and 4 do not hear it, whereas 3 does. If node 6's transmission range is increased (e.g., by boosting its transmission power), the connectivity graph may eventually include either or both the links from 6 to 7 and 4.

Nodes are traversed by *flows* (i.e., distinguishable streams of traffic). Let Q denote the set of all flows. Flow $q\in Q$ is to be routed through a path $P_q\subseteq E$ between its source $s(q)$ and destination $d(q)$. We define the flow's *worst-case end-to-end delay* (*WCD* for short) as the maximum time it takes for a bit of that flow to traverse the whole WMN from its source to the destination, under working conditions. Each flow specifies a *deadline* δ_q, and requests that its WCD be upper bounded by δ_q. At the ingress node, a flow's arrivals are constrained by a *leaky-bucket* shaper, with a *burst* σ_q and a *rate* ρ_q. Packets of each flow are buffered separately at each link. The purpose of this paper is to describe a joint routing and link scheduling scheme that computes a conflict-free schedule which does not violate the required delay bounds whenever it is possible to do so. We first identify the constraints that ensure the conflict-free property, and then move to describing those related to delay feasibility.

The physical interference phenomenon is modeled by means of the widely used *protocol interference* models ([4], [12]). For each edge of the network $e \in E$ we define a conflicting set of edges $\mathcal{I}(e)$ which includes all the edges belonging to E which interfere with e ($\mathcal{I}(e)$ contains e itself); the interference condition is straightforwardly defined as follows:

$$\sum_{i \in \mathcal{I}(e)} x_i(t) \le 1 \text{, if } e \text{ is active in slot } t = 1,2,...,N \text{ ,}$$

where $x_e(t)$ is a binary variable, such that $x_e(t) = 1$ if link $e \in E$ is active in slot t, and 0 otherwise. This means that, if edge e is active in slot t, the associated interfering set $\mathcal{I}(e)$ must contain one active edge only (which is the edge e itself). We translate the interference condition to a *conflict graph* $G_c = (E, C)$, shown in Figure 1, whose nodes are the set of *links* of the connectivity graph and whose edges $C = \{c_1, \ldots, c_r\}$ model the conflicts within the network.

Half-duplex constraints are implicitly accounted for into the interference constraints, links being unidirectional. Hence a set $\mathcal{I}(e)$ can be easily obtained by retrieving the one-hop neighborhood of e in the conflict graph, e.g. for Figure 1 we have $\mathcal{I}(7,8) = \{(4,7),(5,8),(8,7)\}$. Given a conflict graph C, only conflicts between *active links*, i.e. those with a non-null flow, have to be considered. We thus define $C_f \subseteq C$ as the subset of conflicts involving active links:

$$C_f := \{(i,j) \in C : f_i > 0 \text{ and } f_j > 0\},$$

where f_i denotes the flow going through link i.

Following the notation in [8]-[10], we define an *activation offset* π_e for link e, $0 \le \pi_e \le N$, and its *transmission duration* Δ_e. Since time is slotted, both are non-negative integers. Figure 2 shows the above quantities, plus others that will be defined in the following. The assumption that *one* (instead of several) activation of a link in a frame is allowed stems from the fact that, in several technologies (e.g., WiMAX) the link scheduling map is communicated to the various nodes of a WMN in-band: in this case, the shorter the map is, the smaller the overhead is.

The schedule must ensure the *conflict-free* condition: while a link is transmitting, all conflicting links must refrain from transmitting. For any pair of links i and j which are neighboring nodes in C_f we have:

- if j transmits after i, it must wait for i to complete the transmission, i.e. $\pi_i - \pi_j + \Delta_i \le 0$.
- Otherwise, the symmetric inequality holds, i.e. $\pi_j - \pi_i + \Delta_j \le 0$

In order to linearize the combination of the above constraints, we introduce a binary variable o_{ij}, $(i,j) \in C_f$, which is 1 if i transmits after j, 0 otherwise. The left-hand side of the previous constraints can thus be upper bounded by N regardless of the relative transmission order, as π_i and Δ_i belong to $[0, N]$. This completes the formulation of the *conflict-free constraints*, which are necessary and sufficient conditions:

$$\begin{aligned} &\pi_i - \pi_j + \Delta_i \le N \cdot o_{ij} && \forall (i,j) \in C_f \\ &\pi_j - \pi_i + \Delta_j \le N \cdot (1 - o_{ij}) && \forall (i,j) \in C_f \end{aligned} \quad (1)$$

For a schedule to be valid, each link must also complete its transmission within the frame duration, i.e.:

$$\pi_i + \Delta_i \le N \quad \forall i \in E \text{ .} \quad (2)$$

Additional constraints are needed to keep into account the end-to-end delay requirements. During its activation, each link e transmits traffic of all the flows that traverse that link. We can therefore partition the link's Δ_e among them, i.e. $\Delta_e = \sum_{q: e \in P_q} \Delta_e^q$. Δ_e^q is the link activation quota reserved for flow q, which needs *not* be an integer, since when a link e is activated it can switch among backlogged queues regardless of slot boundaries. We assume that backlogged flows traversing e are served in the same (arbitrary) local order, and we call I_e the ordered set of the flow indexes. We assume that each backlogged flow q is served for *no less* than Δ_e^q. If a flow is idle, its service time can be exploited by other backlogged flows at e, as long as the transmission from any flow z starts within at most $\sum_{x \in I_e : x < z} \Delta_e^x$ from π_e. Therefore, flow q has a *guaranteed rate* at link e equal to:

$$R_e^q = W_e \cdot \Delta_e^q / N \text{ .} \quad (3)$$

Since each flow transmits once per frame, a maximum *inter-service time* is guaranteed for that flow, and it is equal to:

$$\theta_e^q = \left(N - \Delta_e^q\right) \cdot T_S \text{,} \quad (4)$$

irrespective of the local ordering at each link. Therefore, each link of a mesh router is a *rate-latency* server [11] for the flows traversing it, with a rate R_e^q and a latency θ_e^q. Accordingly, each flow's WCD is equal to (see [11]):

$$D_q = \begin{cases} \sum_{e \in P_q} \theta_e^q + \sigma_q / R_{\min}^q & \text{if } \rho_q \le R_{\min}^q \\ \infty & \text{otherwise} \end{cases} \text{,} \quad (5)$$

where $R_{\min}^q = \min_{e \in P_q} \{R_e^q\}$. The first addendum in (5) is called *latency delay*, and it is due to link scheduling and arbitration of the flows at the links. The second is called *burst delay*, and it is the time it takes for the flow's burst to be cleared at the minimum guaranteed rate.

Given the traffic, the network topology and the conflict graph, our purpose is to find a joint conflict-free routing and scheduling which is also feasible from a delay point of view. To achieve this, we formulate the *Delay-Aware Routing and Scheduling (DARS) problem* as follows:

$$\begin{array}{lll}
\min \quad V_{\max} & & \\
\text{s.t.}: & & \\
\sum_{e\in E}\theta_e^q + \dfrac{\sigma_q}{R_{\min}^q} - \delta_q \le V_{\max} & \forall q\in Q & (i) \\
(N-\Delta_e^q)\cdot T_S \le \theta_e^q + N\cdot T_S\cdot(1-t_e^q) & \forall e\in E, \forall q\in Q & (ii) \\
R_{\min}^q \le \dfrac{W_e}{N}\cdot\Delta_e^q + (1-t_e^q)\cdot\max_{i\in E}\{W_i\} & \forall e\in E, \forall q\in Q & (iii) \\
\Delta_e^q \ge \dfrac{\rho_q}{W_e}\cdot N\cdot t_e^q & \forall e\in E, \forall q\in Q & (iv) \\
\Delta_e^q \le N\cdot t_e^q & \forall e\in E, \forall q\in Q & (v) \\
\Delta_e \le N\cdot\sum_{q\in Q} t_e^q & \forall e\in E & (vi) \\
\Delta_e \ge \sum_{q\in Q}\Delta_e^q & \forall e\in E & (vii) \\
\pi_i + \Delta_i - \pi_j \le \left(1-o_{ij}\right)\cdot N & \forall i,j\in C_f & (viii) \\
\pi_j + \Delta_j - \pi_i \le o_{ij}\cdot N & \forall i,j\in C_f & (ix) \\
\pi_i + \Delta_i \le N & \forall i\in E & (x) \\
\sum_{e\in OUT(v)} t_e^q - \sum_{e\in IN(v)} t_e^q = \begin{cases} 1 & \text{if } v=s(q) \\ -1 & \text{if } v=d(q) \\ 0 & \text{otherwise} \end{cases} & \forall v\in V, \forall q\in Q & (xi) \\
t_e^q, o_{ij}\in\{0,1\},\ R_{\min}^q, \theta_e^q, \Delta_e^q \ge 0,\ \Delta_e, \pi_e \in Z_0^+ & \forall e\in E, \forall q\in Q &
\end{array}$$

The objective function to be minimized is the maximum *deadline violation* $V_{\max}$, defined as $V_{\max} = \max_{q\in Q}\left\{D_q - \delta_q\right\}$. If the optimum is negative, then the DARS problem has a solution which is feasible from a delay point of view. There are two sets of variables, related to link scheduling (o_{ij}, π_e, Δ_e) and routing (t_e^q) decisions. As for routing, $t_e^q = 1$ iff. flow *q* traverses link *e*. As single-path (as opposed to multipath) routing is assumed, t_e^q are binary. Constraints (*xi*) ensure flow conservation at each node. Constraints (*i-vii*) ensure a delay-aware link scheduling. Specifically, (*i*) represents $D_q - \delta_q$ according to (5) for flow *q*, assuming that its delay is finite. Constrains (*ii-iv*) include at the right hand side terms which depend on $(1-t_e^q)$ and t_e^q. Those terms are computed such that, if $t_e^q = 0$, then the constraints always hold regardless of the value given to Δ_e^q, θ_e^q, $R_{\min}^q$. In other words, those constraints are *inactive* for those links that are not traversed by a flow. On the other hand, when $t_e^q = 1$, (*ii*) sets the latency according to (4), (iii) guarantees that $R_{\min}^q$ is the minimum guaranteed rate among all the links traversed by flow *q*, i.e. $R_{\min}^q = \min_{e:t_e^q=1}\left\{W_e\cdot\Delta_e^q / N\right\}$, and *(iv)* ensures that the activation quota for flow *q* is set according to (3), thus ensuring that the delay is finite. On the other hand, constraints (*v*) and (*vi*) are active when $t_e^q = 0$, when they guarantee that Δ_e^q is forced to zero when flow *q* does not traverse link *e*. Those constraints always hold when $t_e^q = 1$, instead. Constraint (*vii*) relates the activation of a link with the activation quotas of each flow traversing it. Constraints (*viii-x*) mirror (1)-(2), and are thus related to conflict-free scheduling.

Note that, since the routing is specified as part of the model, the latter allows one to account for both *local* traffic, directed from one node to another, and *Internet* traffic, directed from/to an Internet gateway node (i.e., both uplink and downlink). Furthermore, if the WMN has *more* than one gateway node, a straightforward modification of the model allows one to perform *gateway selection*, i.e. to select the gateway through which each flow has to be routed to guarantee the best objective. As shown in Figure 3, all it takes is to add a *virtual* super-gateway node, connected solely to the gateways via mutually non-interfering links of suitable capacity (e.g., T1 or higher), and to select the latter as the source/destination node for *all* the Internet traffic.

The DARS problem is a *Mixed Integer Non-Linear (MINLP)* problem, whose non-linear constraints are convex and for which efficient general purpose MINLP solver (e.g. [13],[14]) exist. The latter can be easily re-formulated as a *quadratic* problem by introducing auxiliary variables, which makes it possible to use the efficient solver CPLEX [13]. Despite the quadratic formulation, the solution time of the above problem is prohibitive for mesh networks of medium to large size. For instance, CPLEX may take days to find the optimum for a 4×4 grid, and cannot solve a 5×5. For this reason, in the next section we present a heuristic approach to solve the DARS problem.

Before moving to the heuristics for the DARS, we justify the need to solve the routing and link scheduling jointly via a simple example. Figure 4 reports a sample 4×4 grid mesh, where four homogeneous flows need be routed from their source (nodes 0-3) to the gateway (node 15). It is $\sigma = 1000$, $\rho = 2000$, $\delta = 30$ for all flows. The link capacity is $W = 9600$ for all links *except* (7,11), whose capacity is 5000. The figure also reports the routes selected by the DARS (the other variables are omitted for ease of reading). A quick glance suffices to convince the reader that these routes are not shortest paths, and it takes only a little more to verify that no shortest-path routing leads to a feasible link scheduling: for instance, if flow 3 were routed along its shortest path 3-7-11-15, then link (7,11) should carry at least 2000 units of rate, i.e. be active for at least 40% of the time. This would leave no more than 60% for conflicting link 11-15 which would then be unable to support flows 1, 2, 3 together. The latter, in fact, require an activation of at least 62.5% on that very link just to keep their WCD *bounded* (since $3\cdot\rho = 0.625\cdot W$), let alone below any pre-specified deadline.

Accepted version of:
P. Cappanera, L. Lenzini, A. Lori, G. Stea, G. Vaglini, "Optimal joint routing and link scheduling for real-time traffic in TDMA Wireless Mesh Networks", Elsevier Computer Networks, 57(11), 5 August 2013, pp. 2301–2312, DOI 10.1016/j.comnet.2012.11.021

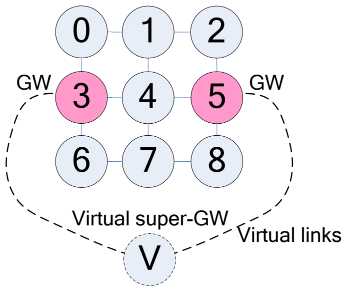

Figure 3. Connectivity graph of a WMN with multiple gateways and gateway selection.

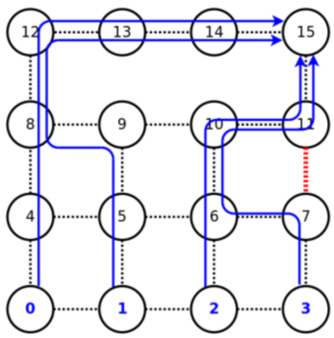

Figure 4. Sample mesh

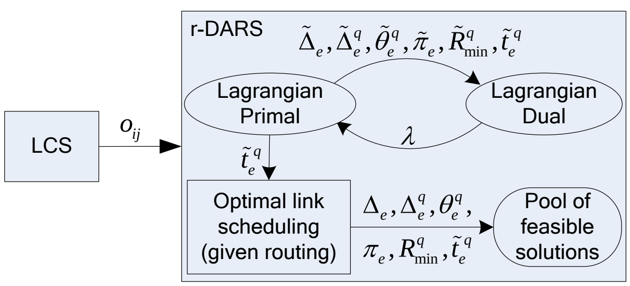

Figure 5. Separate heuristic approach

III. HEURISTIC SOLUTIONS

The high complexity of the DARS problem stems from the high number of binary variables related to conflict (o_{ij}) and routing (t_e^q). Of course, we cannot separate the routing variables without incurring in the problems outlined in the previous example. Therefore, in order to reduce the computation time, we separate the *link conflict serialization (LCS)* from the DARS problem. In other words, we set the o_{ij} variables *offline*, based on the conflict graph, and then solve the reduced DARS, where the o_{ij} are constants. As we will show later on, this allows larger-scale problems to be solved, with a negligible loss of accuracy. To increase the scale further, we also propose a Lagrangian heuristic to solve the *reduced* DARS (r-DARS henceforth) problem suboptimally. We first describe how to solve the LCS, and then we move to the r-DARS. Our solution scheme is detailed in Figure 5.

A. Link Conflict Serialization

Solving the LCS problem consists in setting the o_{ij} variables, i.e. *directing* the edges in the conflict graph, which in turn translates to *serializing* conflicting links within the frame. In fact, all the links belonging to the same clique in the conflict graph – e.g., (0,1), (1,4) in Figure 1 – cannot be activated in parallel, hence have to be serialized. Assuming for instance one-hop interference (which is not a requisite, in any case), a link may belong to up to two cliques (i.e., those of either ends). For instance, (0,1) also belongs to a 3-clique with (3,0) and (0,3), and to a 2-clique with (1,4). We remark that one-hop interference is not a mandatory assumption in our model. The objective to be pursued by the LCS is thus to *minimize the maximum path length* in the resulting directed conflict graph[1], i.e. to minimize the maximum number of serialized links. On one hand, this allows larger activations to be given to the links in the maximum-length path. More importantly, however, this allows greater flexibility in allocating activation time, once routing and link scheduling determine the load that flows impose on each link.

The LCS can be solved by employing a general K-coloring method [32]. The K-coloring is exponential in the number of vertices. However, it can be solved up to scales much larger than the ones we are dealing with, and efficient methods – e.g., based on column generation [36] – can be exploited to solve the problem at larger scales.

Thus the LCS can be solved optimally, given the conflict graph. Therefore, as traffic changes, a new routing and link scheduling can be computed without modifying the conflict serialization. The negative side of solving the LCS without taking traffic into account is that a possibly *short* path in the conflict graph (i.e., one with few links) may end up carrying a large amount of traffic because of routing, and hence become critical. Nevertheless, since routing decisions are taken afterwards in the r-DARS, flows would be routed around such critical paths as a consequence of routing decisions.

B. Lagrangian heuristic

The r-DARS is still a complex problem. While it can be solved in a matter of seconds in a 4×4 grid, it takes hours to solve it on a 5×5 grid. Therefore, we propose a heuristic scheme to solve it. The design of the heuristic should pursue the following two objectives: (i) exploit the very structure of the r-DARS problem, where two strictly interrelated decisions (i.e., routing and scheduling) are to be made; (ii) provide a quantitative metric that is able to measure the quality of the so-

[1] *Paths* in the conflict graph are obviously different from those in the *connectivity* graph.

Accepted version of:
P. Cappanera, L. Lenzini, A. Lori, G. Stea, G. Vaglini, "Optimal joint routing and link scheduling for real-time traffic in TDMA Wireless Mesh Networks", Elsevier Computer Networks, 57(11), 5 August 2013, pp. 2301–2312, DOI 10.1016/j.comnet.2012.11.021

lutions. In fact, the goodness of a heuristic solution is usually measured in terms of relative gap between its value and the optimum solution value (if available). Since, in our settings, the optimal solution of r-DARS is affordable for medium-sized instances only, we have to make do with lower bounds to the optimal value in larger instances. Our choice is to propose a *Lagrangian relaxation-based* heuristic, which is a mathematical tool widely acknowledged in the literature as a means to get this twofold advantage. Specifically, in our problem a Lagrangian heuristic allows to: (i) decompose the r-DARS, gaining in efficiency and/or scale; (ii) compute a lower bound that is demonstrably not worse than the straightforward bound given by the relaxation of the integer constraints (the so-called *continuous relaxation*), in addition to giving an upper bound. We first explain how to obtain a Lagrangian relaxation, and then show how the heuristic is built upon the latter.

The r-DARS has two blocks of variables: the link scheduling variables, involved in constraints (*i*), (*vii-x*) and the routing variables in constraint (*xi*). In addition, a set of *coupling constraints*, i.e. (*ii-vi*), collate link scheduling and routing decisions. In the absence of the latter, r-DARS could be decomposed in two subproblems: a link scheduling problem and a routing problem respectively. Hence we perform a Lagrangian relaxation with respect to the coupling constraints: rather than eliminating the complicating constraints, the latter are dualized by inserting them in the objective function and associating a non-negative Lagrangian multiplier λ_i with each of them. For a given setting of λ, the *Lagrangian primal problem* to be solved is the following:

$$\begin{aligned}\varphi(\lambda) = &\min_{s.t.\,(i),\ (vii\text{-}x)} \left\{ V_{\max} + s(\lambda;\Delta_e,\Delta_e^q,\theta_e^q,\pi_e,R_{\min}^q) \right\} \\ &+ \min_{s.t.(xi)} \left\{ r(\lambda;t_e^q) \right\}\end{aligned}, \quad (6)$$

where $s(\lambda;\Delta_e,\Delta_e^q,\theta_e^q,\pi_e,R_{\min}^q)$ and $r(\lambda;t_e^q)$ are linear cost functions depending on the Lagrangian multipliers (*updated Lagrangian costs*). The Lagrangian multiplier λ_i plays two roles: i) it penalizes the variables for which the relaxed *i*-th constraint is violated by adding a positive term to the original objective function, and ii) it favors solutions for which the relaxed *i*-th constraint is satisfied, by adding a negative term to the objective function. Function $\varphi()$ is *separable*: for a given value of λ, solving the Lagrangian primal implies solving separately a scheduling problem and a routing problem, which is considerably faster than solving them jointly. Yet this scheme keeps routing and scheduling together through the multipliers, hence retaining the benefits of a joint approach. The solution thus computed is a *lower bound* on the optimum of the r-DARS for each choice of the Lagrangian multipliers. It is thus necessary to compute the *best* lower bound among the possible choices of λ, i.e., to solve the *Lagrangian dual*:

$$\max_{\lambda \ge 0} \left\{ \varphi(\lambda) \right\}. \quad (7)$$

The Lagrangian dual is solved via an iterative algorithm which alternates between a *primal phase*, where routing and scheduling problems are solved separately for a given λ, and a *dual phase*, where information gathered in the primal phase (i.e., the solution of the two problems and the violation of the coupling constraints) are collected and mixed together to update the value of λ accordingly..

It is also evident that the routing variables play a key role in this Lagrangian scheme. For a given choice of the Lagrangian multipliers, the routing problem results in a Minimum Cost Multicommodity Flow problem, where a path has to be computed for each source-destination flow so as to guarantee flow balance constraints as well as global capacity constraints on the links. The costs, to be minimized, depend on the Lagrangian multipliers as shown in (6). The routing problem is solved via CPLEX. At each iteration, once the routing problem has been solved (i.e., a path for each flow is known), an attempt to construct a feasible solution can be done by solving a scheduling problem in cascade (*optimal link scheduling* in Figure 5). This step entails solving a *Mixed Integer Non-Linear* problem, whose non-linear constraints are convex. If a feasible link scheduling is computed on a given routing, then the solution verifies *all* the constraints, and is thus admissible for the r-DARS problem (although not necessarily optimal), hence it is an upper bound on the optimum. As the Lagrangian scheme is iterated, possibly many feasible solutions are computed this way and stored in a pool. When the Lagrangian dual is solved:

a) the best feasible solution in the pool is returned.
b) the best lower bound is given.

Note that, even though routing and link scheduling are decided in two separate modules in Figure 5 (i.e., the Lagrangian primal and the optimal link scheduling), the fact that the Lagrangian scheme iterates between the primal and dual, computing bounds on the activation variables, implies that routing decisions are affected by scheduling decisions and vice-versa, which makes the approach *joint* in all respects.

A solution approach like this belongs to the *Lagrangian heuristics* family ([31]). In our approach the Lagrangian dual is solved via a *bundle* type method ([29]-[30]). The latter is an iterative ascent algorithm where both the ascent direction and the step along that direction needed to update the Lagrangian multipliers at each iteration are chosen based on information collected during the previous iterations. A bundle method differs from a *subgradient* approach, which is a classical method used to solve the Lagrangian dual, where the multipliers are

Accepted version of:
P. Cappanera, L. Lenzini, A. Lori, G. Stea, G. Vaglini, "Optimal joint routing and link scheduling for real-time traffic in TDMA Wireless Mesh Networks", Elsevier Computer Networks, 57(11), 5 August 2013, pp. 2301–2312, DOI 10.1016/j.comnet.2012.11.021

updated according to the information collected in the *last* iteration only. Being based on a more global perspective of the problem, a bundle algorithm is generally more efficient..

IV. PERFORMANCE EVALUATION

The contribution of this section is twofold. First, we evaluate the performance of our heuristic approach to solve the DARS problem, in terms of optimality and complexity. Second, we exploit it to infer structural properties of the WMN, i.e. optimal placement of one or more Internet gateway nodes and analysis of the schedulability as a function of the transmission radius. We present the above contributions in separate subsections.

A. Evaluation of the heuristic approach

As for the first objective, we make simulations on a grid of varying diameter, up to 7×7 nodes. All links have a capacity equal to 9600, and the gateway is located in one corner. We assume that each link interferes only with those that are one hop away, and set the conflict graph accordingly[2]. One flow is originated at each node, and is to be routed to the gateway. Instances are solved using an Intel Core 2 Duo CPU, 2.33GHz using IBM ILOG CPLEX 12.1

As for optimality, we compare the *optimal DARS solutions*, where available (up to a 4×4 grid) and those computed with the heuristic *LCS+r-DARS*. In this last approach, the r-DARS is solved both optimally and via the Lagrangian heuristic. For each test set, we evaluate the objective on a set of 30 randomly generated instances, with heterogeneous flow requirements: rates and bursts are generated uniformly between $[0,9600/(2\cdot|Q|)]$ and $[0,1000]$, while the deadlines are set to either 60 or 90. Frames have 100 slots. We first show that separating the LCS and the r-DARS yields accurate results. Figure 6 shows the relative gap with respect to the DARS optimum in a 4×4 grid. The figure clearly shows that the suboptimal solutions of the two schemes are within few percentage points to the optimum.

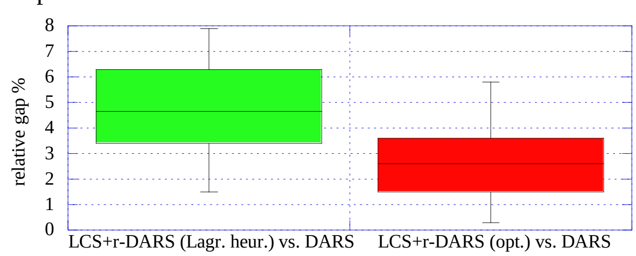

Figure 6. Accuracy comparison of the heuristic schemes

[2] Other choices, e.g. two-hop interference, can also be accommodated in our model.

However, solving the r-DARS optimally is time consuming: already with 5×5 grids, we could not find instances this took less than 8000s. Instead, the Lagrangian heuristic is considerably faster. Figure 7 reports a box plot of the solution times of 30 instances of grids, from 4×4 to 7×7. The figure shows that routing plans can be done in a few hours for grids up to 7×7, which is quite a large dimension for a WMN.

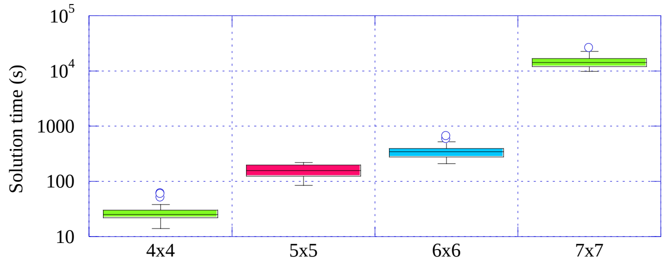

Figure 7. Solution time for the LCS+r-DARS, using the Lagrangian heuristic

Next, we show the benefits of having a *joint* routing and scheduling, by comparing it to a *cascading* approach, where routing decisions are taken first, oblivious of link scheduling. In the latter, we use a *capacitated multicommodity flow* (CMF) routing, where each flow q requires a capacity equal to its rate ρ_q, and the routing that minimizes the overall number of traversed links is chosen, keeping into account the capacity constraints. The CMF sets the t_e^q variables, and then the link scheduling is solved optimally given the routing, as in [9]. In the *joint* approach, we use LCS+r-DARS, with the latter solved through the Lagrangian heuristic. Figure 8 shows the relative gap between the cascading and the joint approaches for two sets of instances of a 6×6 grid: for the first set rates and burst are again generated uniformly between $[0,9600/(2\cdot|Q|)]$ and $[0,1000]$, for the second one the rates are generated between $[0,9600/(1.2\cdot|Q|)]$; this leads to instances where the WMN is highly congested, with the links close to the gateway approaching the saturation point. For the first set a joint approach (although solved suboptimally) always performs 10%-15% better in terms of objective function, *despite* the fact that both subproblems are solved optimally in the cascading approach. For the second set the gap grows to 20%. However, the cascading approach fails to compute a feasible link schedule in as many as 37% of the instances, whereas our joint approach solves them all.

Then, we show how schedulability of a set of flows changes with their rate and burst. Figures 9-11 show the maximum violation as a function of the burst and rate of the flows. Figures 9 and 10 show results for a burst value of 1000 against a rate from 50 to 300 on a 5×5 and 6×6 grid respectively. Figure 11 reports results for a burst size ranging from 0 to 2000 and a rate of 150. In the above figures, the (unfeasible) solution of the

Accepted version of:
P. Cappanera, L. Lenzini, A. Lori, G. Stea, G. Vaglini, "Optimal joint routing and link scheduling for real-time traffic in TDMA Wireless Mesh Networks", Elsevier Computer Networks, 57(11), 5 August 2013, pp. 2301–2312, DOI 10.1016/j.comnet.2012.11.021

continuous relaxation of the r-DARS problem is shown for comparison. The latter is a lower bound on the optimum, and its purpose is to show that – despite we cannot compute the optimum DARS solution – both the r-DARS optimum and its heuristic approximation are quite close to the DARS lower bound, hence to the DARS optimum itself. Note that in the continuous relaxation routing variables are not integer. In this case, constraints (ii-vi) in the DARS model have no physical counterparts. This justifies the fact that the lower bound is hardly affected by the rates and bursts in Figures 9-11.

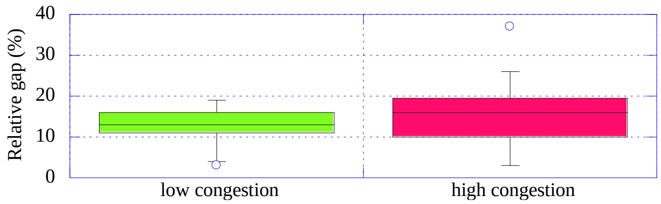

Figure 8. Relative gap between the cascading and the joint approach (the latter solved through the Lagrangian heuristic) on a 6x6 grid WMN

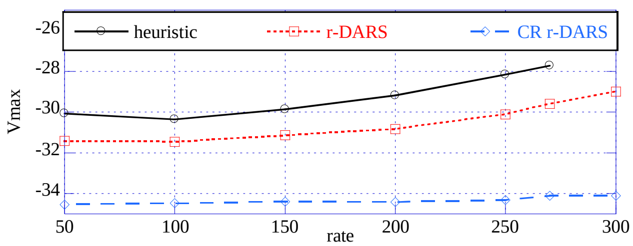

Figure 9. Maximum violation as a function of the rate for a 5x5 grid topology

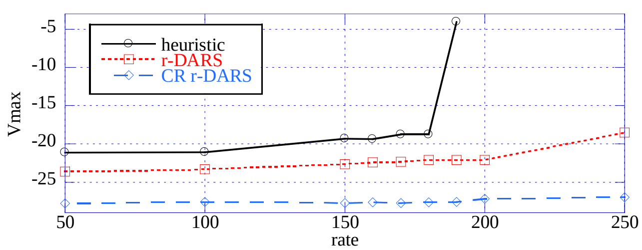

Figure 10. Maximum violation as a function of the rate for a 6x6 grid topology

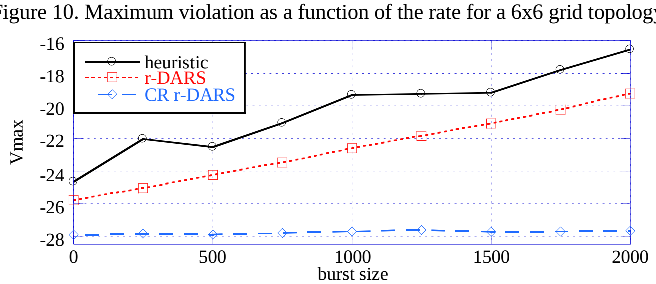

Figure 11. Maximum violation as a function of the burst size

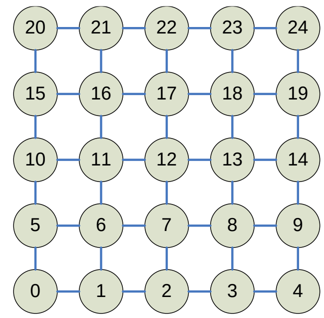

Figure 12. The test-case 5x5 WMN

B. *Case study: optimal gateway placement*

We now show how to exploit our solution scheme to infer properties which are useful from a network engineer perspective. More specifically, we discuss optimal gateway placement in both single-gateway and multi-gateway WMNs. We take as an example a 5×5 grid mesh, shown in Figure 12. The latter has 25 nodes, 80 links and 416 conflicts. We initially place a single gateway and homogeneous traffic, one flow from each node to the gateway. For obvious reasons of symmetry, we only move the gateway toward one border and corner of the WMN. Figure 13 shows V_{max} as a function of the rate when a single gateway is placed at various nodes, from the center to the corner, for a burst equal to 1000 and a deadline of 60. The figure shows that V_{max} is minimized when the gateway lies in the center. The result makes sense since a central gateway minimizes the length of the *longest* path as well, which are the ones likely to contribute to V_{max}. Figure 14 further clarifies that a larger V_{max} is obtained in conjunction with a higher resource expenditure, its vertical axis reporting the sum of the allocated capacity on all the slots of the schedule. Note that it is not possible to obtain a feasible schedule with $\rho=350$ when the gateway is placed in the corner.

We repeated the evaluation with random flows, whose parameters are the same as in the previous section. The results, shown in Figure 15, show that the distribution of V_{max} moves to the right as we move the gateway from the centre to one corner.

Finally, we compared the single-gateway scenario to one where the WMN has two gateway nodes. Figure 16 shows both V_{max} (left vertical axis) and the allocated capacity (right vertical axis) as a function of the placement of the gateways. The most favorable single-gateway scenario is reported on the left for comparison. All data are related to a homogeneous traffic scenario, with one flow from each non-gateway node whose characteristics are $\rho=100$, $\sigma=1000$ and $\delta=90$. Note that the two-gateway scenarios have one flow less than the single-gateway scenario, as gateways send no traffic themselves. The figure shows that the more far apart the two gateways are, the worse V_{max} is, and the higher (in general) is the allocated ca-

Accepted version of:
P. Cappanera, L. Lenzini, A. Lori, G. Stea, G. Vaglini, "Optimal joint routing and link scheduling for real-time traffic in TDMA Wireless Mesh Networks", Elsevier Computer Networks, 57(11), 5 August 2013, pp. 2301–2312, DOI 10.1016/j.comnet.2012.11.021

pacity. However, it also shows that the only result that can be achieved by putting two gateways is to improve V_{max} marginally, at the price of a 27% increase in the allocated capacity. Within the limit of the considered scenarios, this suggests that a *single* gateway, placed at the center, is the optimal solution for a WMN of this topology and traffic.

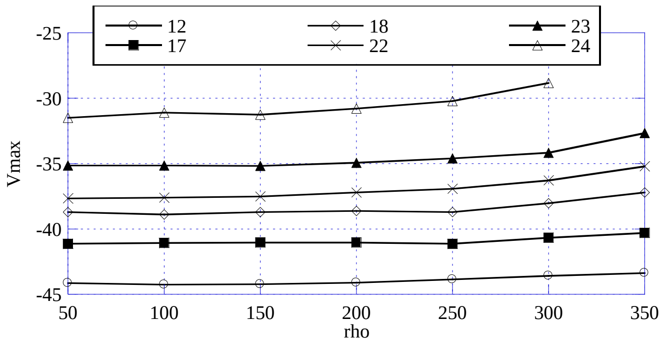

Figure 13 – Vmax as a function of the rate for various gateway placements – homogeneous traffic

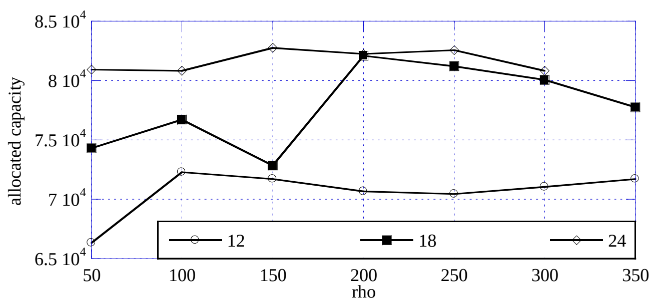

Figure 14. Allocated capacity as a function of the rate for various gateway placements – homogeneous traffic

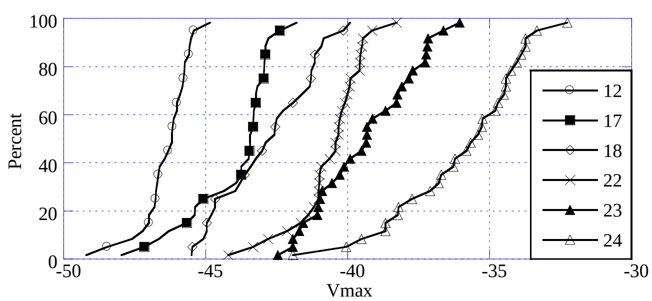

Figure 15. Distribution of Vmax over 30 random instances with different placements of the gateway node

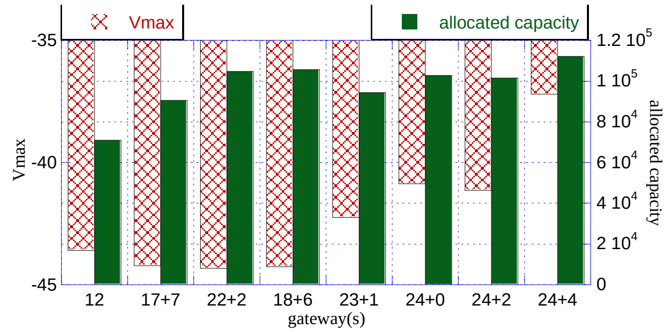

Figure 16. Vmax and allocated capacity for a single-gateway and two-gateway scenario

C. *Case-study: schedulability as a function of the transmission radius*

We now show how the schedulability of a set of flows is affected by the transmission radius. We consider a WMN and increase the transmission range of each node, so that distant nodes progressively get connected. On one hand, this increases the number of conflict, making link scheduling harder. On the other hand, the number of hops in a path is bound to decrease, which counterbalances the above effect.

The WMN we consider is loosely based on the one of the TFA project at Rice University, [39], and is shown in Figure 17. We deploy the 21 nodes in the same positions as in [39], and assume that each node is equipped with one omnidirectional antenna[3]. We set the transmission range of each node so that the WMN is fully connected (the resulting logical connectivity graph is in fact the one shown in Figure 17) and vary it by multiplying each range by a constant factor $M \in [1;1.6]$. The capacity of the links is constant and equal to 5000. We setup the flows as shown in Table 1. Figure 18 reports the number of edges and conflicts as a function of M. Both are increasing, alternating plateaus and steps, the latter occurring when the transmission range reaches some critical inter-node distance. Furthermore, the number of conflicts increases slightly faster than the number of edges, which is also expectable, given that each new link to a destination conflicts with potentially many links. Note that - already with $M = 1$ - the number of both edges and conflicts is higher than those of the previous case study of Figure 12, which has 80 edges and 416 conflicts. Thus, this case study is significantly more complex than the former, despite having fewer nodes.

Figure 19 shows V_{max} as a function of M, for both the heuristic and the optimal solution of the r-DARS. The figure shows that the shortening of paths prevails over the increase in the

[3] In [39], some nodes are also equipped with directional antennas to gateway nodes. We do not include these links, which are less interesting from a link scheduling perspective.

Accepted version of:
P. Cappanera, L. Lenzini, A. Lori, G. Stea, G. Vaglini, "Optimal joint routing and link scheduling for real-time traffic in TDMA Wireless Mesh Networks", Elsevier Computer Networks, 57(11), 5 August 2013, pp. 2301–2312, DOI 10.1016/j.comnet.2012.11.021

number of conflicts, hence V_{max} decreases with M. Furthermore, the heuristic gets closer to the optimum as M increases.

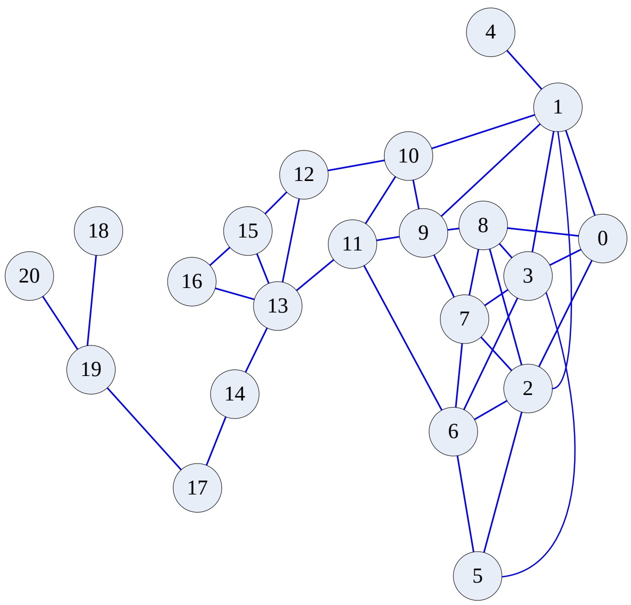

Figure 17. The case-study WMN. Edges represent links obtained in the base case $M = 1$

Table 1. Flows for the case study

Flow #	source	dest	σ	ρ	δ
1	20	4	1000	800	10
2	5	18	1000	1000	10
3	17	0	1000	800	10
4	16	5	1000	800	10
5	5	4	1000	600	10
6	1	16	1000	800	10

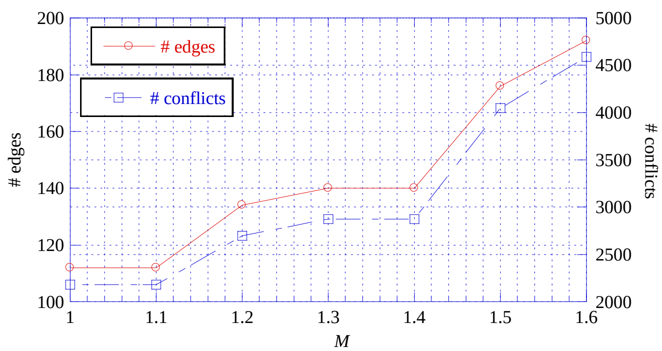

Figure 18. Number of edges and conflicts as a function of the transmission range.

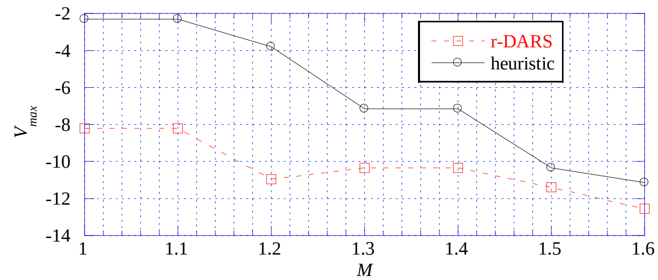

Figure 19. V_{max} as a function of the transmission range

V. Related Work

In this section we review some of the related works on routing and link scheduling in WMNs. As the literature on the subject is abundant, we narrow down the scope to those that are more germane to our work, leaving out anything connected with multi-radio systems (where the *channel assignment* problem is the most prominent issue) and/or not dealing with performance bounds. As already stated, no work that we are aware of (save our previous work on the same topic, [8]-[10]) considered schedulability in WMNs with: i) VBR traffic, and ii) arbitrary deadline constraints. Most of the link scheduling approaches fall into either of the following categories:

1. *rate-oriented* algorithms, that either provide flows with a *minimum guaranteed rate* (e.g. [15]-[18]), or optimize the total throughput (e.g. [19]-[22]). Guaranteeing a minimum rate no smaller than the flow's rate – e.g. by (5) – is a necessary condition for WCDs to be *finite*, but does not automatically make them smaller than a pre-specified deadline. In fact, by renouncing over-allocating rates, these schemes often compute schedules with unfeasibly large WCDs.
2. *TDMA delay-oriented* algorithms, that either minimize (e.g. [25]-[26]) or try to guarantee a *maximum TDMA delay* (e.g. [23]-[24]). The latter is the sum of TDMA waiting times at every hop, i.e. the time it takes for a packet to travel from the source to the destination, assuming that it is never queued behind other packets. As queuing is a component (and often the dominant one) of the end-to-end delay, especially with VBR traffic, there is no guarantee that such algorithms can actually find a deadline-feasible schedule if there exists one. We show this later on, using [26] as a comparison.

Within the second category, [25] considers both CBR (voice) and VBR (video) flows, however assuming that VBR sources can be described as stationary, ergodic and independent processes with known statistics, so as to characterize them as equivalent CBR sources. In this work, we deliberately omit this kind of assumptions, sticking instead to more practical σ, ρ characterizations, which can be conveyed to the network using standard signaling protocols such as RSVP). In [26], a WMN is modeled as a stop-and-go system. A min-max problem on the round-trip TDMA delay introduced by the scheduling in a sink-tree network is formulated and optimally solved. To reinforce the point that minimizing the TDMA delay is not the same thing as computing deadline-constraint schedules, we compare our schedules with the optimal ones derived from [26] in a simple sink-tree network (i.e., one where routing is not an issue). In that work, the activation of each link is computed based on the *rate* of the flows traversing it, and activations are serialized so as to minimize the maximum TDMA delay. Consider a WMN of 15 nodes arranged in a binary tree, with homogeneous traffic and 20 uplink flows originating at each node. Fix $\delta = 20$, $\rho = 300$, and let the burst of the flows vary as $0 \le \sigma \le 4500$. We plot the value for $V_{\max}$ obtained by: i) optimally solving the link scheduling according to the DARS, and ii) using the optimal solutions given by [26] in the same settings. As Figure 20, shows, the above traffic cannot be scheduled for bursts larger than 500 according to [26], whereas it is perfectly schedulable in our framework. This is because [26] optimizes *only* conflict orientations (o_{ij}) and activation instants (π_e), neglecting the activation *durations* (Δ_e, Δ_e^q), i.e. renouncing trading *rate* for *delay*. Our work instead explores the other extreme of the rate-delay trade-off by allocating resources based on the requested deadlines.

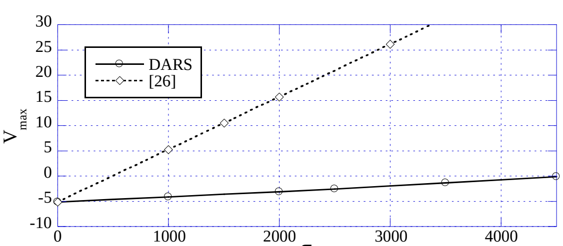

Figure 20. Comparison between optimizing on $V_{\max}$ and minimizing the maximum TDMA delay

Some works not falling into either of the above categories are also relevant, as they provide frameworks for computing bounds on the WCD *a posteriori*, after routing and link scheduling have been planned. In [27] authors define the *odd/even link activation* and routing framework, and employ internal scheduling policies at each link so that the bound on the WCD along a path is roughly double the one obtained in a wired network of the same topology. Authors of [28] show that using throughput-optimal link scheduling and Coordinated-EDF to schedule packets within each link, rate-proportional delay bounds with small additive constants are achieved. Our goal is instead to have pre-specified, *arbitrary* deadlines met through link scheduling.

Finally, some works (e.g., [37]-[38]) consider placing one or more gateways subject to QoS constraint. However, they use *additive*, per-link delays in their computation, which hold regardless of the traffic traversing them.

VI. Conclusions and Future Work

In this work we have analyzed Delay-Aware Routing and Scheduling (DARS) problem for WMNs. We have formulated the problem as an optimization problem, which is however too complex to solve optimally already at relatively small scales (e.g., a 4x4 grid WMN). We have devised a heuristic, based on i) extrapolating the link conflict serialization from the rest of the DARS problem, and ii) solving the reduced DARS problem using a Lagrangian heuristic, which allows one to reap the benefits of a joint routing and scheduling approach, without paying the price of the added model complexity. Our results show that the heuristic scheme is fast and accurate, allowing a network administrator to provision a WMN of several tens of nodes so as to meet pre-specified delay guarantees for real-time traffic. Furthermore, we have used the above technique to provide insight into structural properties of WMNs: for instance, we have identified guidelines for the optimal placing of gateways in the WMN, and studied the schedulability when the transmission range of the nodes varies.

This is the first work having deadlines as constraints, despite the abundant literature on joint routing and scheduling. Future work, which is actively being pursued at the time of writing, will include considering multipath routing, i.e. allowing a traffic flow to be split among several paths in order to balance link utilization.

References

[1] I. F. Akyildiz, X. Wang, W. Wang: "Wireless mesh networks: a survey", *Elsevier Computer Networks*, 47:(4), pp.445-487, 15 March 2005.

[2] R. Nelson, L. Kleinrock. "Spatial TDMA: A collision-free multihop channel access protocol", *IEEE Transactions on Communications*, 33(9), pp. 934-944, 1985.

[3] M. Alicherry, R. Bhatia, L.E. Li. "Joint channel assignment and routing for throughput optimization in multiradio wireless mesh networks", *IEEE JSAC*, 24(11), pp. 1960-1971, 2006.

[4] M.S. Kodialam, T. Nandagopal. "Characterizing achievable rates in multi-hop wireless networks: the joint routing and scheduling problem", in *Proc. ACM MOBICOM'03*, San Diego, CA, US, pp. 42-54, 2003.

[5] V. S. Anil Kumar *et al.* "Algorithmic aspects of capacity in wireless networks", *SIGMETRICS Performance Evaluation Review*, 33(1), pp. 133-144, 2005.

[6] T. Moscibroda, R. Wattenhofer. "Coloring unstructured radio networks", in *Proc. ACM SPAA'05*, New York, US, pp. 39-48, 2005.

[7] S. Ramanathan. "A unified framework and algorithm for channel assignment in wireless networks", *Wireless Networks*, 5(2), pp. 81-94, 1999.

[8] P. Cappanera, L. Lenzini, A. Lori, G. Stea, G. Vaglini, "Link Scheduling with End-to-end Delay Constraints in Wireless Mesh Networks", in *Proc. of IEEE WoWMoM 2009*, Kos, Greece, June 15-19, 2009.

[9] P. Cappanera, L. Lenzini, A. Lori, G. Stea, G. Vaglini, "Optimal Link Scheduling for Real-time Traffic in Wireless Mesh Networks in both Per-flow and Per-path Frameworks", in *Proc. of WoWMoM 2010*, Montreal, CA, June 2010

[10] P. Cappanera, L. Lenzini, A. Lori, G. Stea, G. Vaglini, "Efficient Link Scheduling for Online Admission Control of Real-time Traffic in Wireless Mesh Networks", Elsevier Computer Communications, 34(8), pp. 922-934, 2011

[11] J.-Y. Le Boudec and P. Thiran. *Network Calculus*. Springer Verlag, LNCS 2050, New York, NY, 2001.

[12] P. Gupta, P.R. Kumar. "The capacity of wireless networks", *IEEE Transactions on Information Theory*, 46(2), pp. 388-404, 2000.

[13] ILOG CPLEX software, http://www.ilog.com

[14] P. Bonami and J. Lee, BONMIN Users' Manual, http://projects.coin-or.org/Bonmin

[15] R. Bhatia M. Kodialam, "On Power Efficient Communication over Multi-hop Wireless Networks: Joint Routing, Scheduling and Power Control", in *Proc. of INFOCOM 2004,* pp. 1457-1466, Hong Kong, March 7-11, 2004.

[16] J. Kazemitabar, V. Tabatabaee, H. Jafarkhani "Global Optimal Routing, Scheduling and Power Control for Multi-hop Wireless Networks with Interference", in *Coping with Interference in Wireless Networks*, Springer Signals and Communication Technology, 2011, 59-77, DOI: 10.1007/978-90-481-9990-7_4

[17] R. Cruz and A. Sanathanam, "Optimal Routing, Link Scheduling and Power Control in Multi-hop Wireless Networks", in *Proc. of INFOCOM 2003*, S. Francisco, US, March 30-Apr 3, 2003.

[18] J. El-Najjar, C. Assi, B. Jaumard, "Joint Routing and Scheduling in WiMAX-based mesh networks: A Column Generation Approach", in *Proc. of WoWMoM 2009*, Kos, Greece, June 2009.

[19] Y. Wang, W. Wang, X.-Y. Li, W.-Z. Song, "Interference-Aware Joint Routing and TDMA Link Scheduling for Static Wireless Networks, IEEE Trans. on Parallel and Distributed Systems, 19(12): 1709-1725, Dec. 2008.

[20] J. Crichigno, M. Y. Wu , J. Khoury, W. Shu, "A Joint Routing and Scheduling Scheme for Wireless Networks with Multi-packet Reception and Directional Antennas", in *Proc. of WoWMoM 2009*, Kos, Greece, June 2009.

[21] K. Jain, J. Padhye, V. Padmanabhan, L. Qiu, "Impact of interference on multi-hop wireless network performance", in *Proc. of ACM MOBICOM* 2003, San Diego, CA, Sep. 14-19 2003.

[22] O. Goussevskaia, R. Wattenhofer, M. M. Halldorsson, E. Welzl, "Capacity of Arbitrary Wireless Networks", in *Proc. of INFOCOM 2009*, Rio de Janeiro, BR, Apr 19-25, 2009.

[23] M. Leoncini, P. Santi, P. Valente "An STDMA-Based Framework for QoS Provisioning in Wireless Mesh Networks", in *Proc. of MASS 2008*, Atlanta, Georgia, Sep. 29-Oct. 2, 2008.

[24] J. Zou, D. Zhao "Connection-Based Scheduling for Supporting Real-Time Traffic in Wireless Mesh Networks", IEEE Trans. on Wireless Communications, 8(3):1182-1187, 2009.

[25] H. Shetiya, V. Sharma: "Algorithms for Routing and Centralized Scheduling to Provide QoS in IEEE 802.16 Mesh Networks", in *Proc. of WMuNeP'05*, Montreal, CA, Oct. 13, 2005.

[26] P. Djukic, S. Valaee. "TDMA delay aware link scheduling for multi-hop wireless networks", IEEE/ACM Trans. on Networking, 17(3), pp. 870-883, 2009.

[27] G. Narlikar, G. Wilfong, L. Zhang, "Designing Multihop Wireless Backhaul Networks with Delay Guarantees", Springer Wireless Networks, Aug. 2008.

[28] P. Jayachandran, M. Andrews, "Minimizing end-to-end Delay in Wireless Networks Using a Coordinated EDF Schedule", *in Proc. of INFOCOM 2010*, San Diego, CA, March 15-19, 2010.

[29] A. Frangioni "Generalized Bundle Methods" SIAM Journal on Optimization 13(1), pp. 117 - 156, 2002.

[30] A. Frangioni, G. Gallo "A Bundle Type Dual-Ascent Approach to Linear Multicommodity Min Cost Flow Problems" INFORMS Journal On Computing 11(4), pp. 370 - 393, 1999.

[31] C. Barnhart. Dual Ascent Methods for Large-Scale Multicommodity Flow Problems. Naval Research Logistics, 40, pp. 305–324, 1993.

[32] B. Gendrou, A. Herz, P. S. Louis, "On edge orienting methods for graph coloring", Journal of Combinatorial Optimization, 13(2), pp. 163-178, 2007.

[33] H. Li, Y. Cheng, C. Zhou, W. Zhuang, "Minimizing end-to-end delay: a novel routing metric for multi-radio wireless mesh networks", in *Proc. of INFOCOM 2009*, 46-54, Rio de Janeiro, BR, Apr 19-25, 2009.

[34] M. Cesana, L. Fratta, M. Gerla, E. Giordano, G. Pau, "C-VET the UCLA Campus vehicular testbed: integration of VANET and Mesh networks", in *Proc. of European Wireless 2010*, Lucca, Italy, April 2010.

[35] A. Kandhalu, A. Rowe, R. Rajkumar, C. Huang, C.-C. Yeh, "Real-time video surveillance over IEEE 802.11 mesh networks", in *Proc. of RTAS '09*, S. Francisco, US, April 13 - 16, 2009.

[36] A. Mehrotra, M. A. Trick, "A Column Generation Approach For Graph Coloring", INFORMS Journal of Computing, 8 (4), pp. 344-354, 1996.

[37] Y. Drabu, H. Peyravi, "Gateway Placement with QoS Constraints in Wireless Mesh Networks," in *Proc. of ICN 2008*, pp.46-51, Cancun, Mexico, 13-18 Apr. 2008.

[38] B. Aoun, R. Boutaba, Y. Iraqi, G. Kenward, "Gateway Placement Optimization in Wireless Mesh Networks With QoS Constraints", IEEE JSAC 24(11):2127-2136, Oct. 2006.

[39] TFA project, Rice University, http://tfa.rice.edu/.