Evaluation of a Potential Energy Methodology for Joint Routing and Scheduling in Wireless Mesh Networks

Leonardo Badia
IMT Lucca Institute for Advanced Studies, Piazza S. Ponziano 6, 55100 Lucca, Italy
email: l.badia@imtlucca.it

Abstract—In wireless mesh networks, joint optimization of routing and link scheduling within a time-division multiplexing approach is commonly sought to provide end users with high data rates. However, the strategies proposed to this end usually proceed by means of complex optimization models, which also often rely on oversimplified assumptions, especially for what concerns wireless interference. In the present paper, we draw a novel general framework to perform joint routing and scheduling avoiding these limitations. We evaluate sequences of Link Activation Modes, i.e., sets of transmissions which can be performed simultaneously, and we introduce the concept of potential energy of a mesh network, thanks to which we outline efficient selection of Link Allocation Modes in order to jointly solve routing and scheduling. A heuristic strategy derived within this framework is numerically evaluated by means of simulation and is shown to achieve very good performance, obtained with extremely low computational complexity.

I. INTRODUCTION

Joint routing and scheduling (JRS) represents a very interesting challenge for wireless mesh networks [1]–[3]. Specifically, we focus in this paper on a network with centralized control determining transmission activities over wireless links in a Time Division Multiple Access (TDMA) fashion [2], [4]. In such a case routing and scheduling performed separately often fail to guarantee good performance [5]. For this reason, a Linear Programming (LP) framework [1] can be issued for the joint optimization of both routing and link scheduling. However, the resulting complexity of this cross-layer LP is usually high.

Many optimization strategies rely on specific assumptions about the radio interference. Since most descriptions take a high level perspective, additional approximations are therefore introduced, according to the accuracy by which the interference is captured. In particular, the so-called protocol interference model is used, as defined in [6]. In this way, the interference is represented as a compatibility relationship between links, which is often modeled through a conflict graph [3]. Even though this methodology allows to obtain interference-free transmissions through simple graph coloring algorithms, it also introduces approximations in the fact that interference is not a binary relationship. Moreover, this technique is no longer applicable if a different interference model is used. For these reasons, we do not rely on these specific assumptions about wireless interference.

We emphasize that we do not seek to reduce the optimization search space, but rather to decouple the constraints represented by interference conditions and traffic delivery, which is a more efficient reduction of the problem complexity. This is realized by working on what we call the Link Allocation Modes (LAMs) [7], i.e., set of logical transmissions which can be simultaneously performed according to the constraints related to physical aspects, such as the wireless interference characterization.

Moreover, we present in this paper a novel strategy to sequentially allocate feasible LAMs. We analyze LAMs with an original approach which, to some extent, recalls the management of water flows and drainage systems [8]. We develop a framework where the potential energy of the network is derived, mimicking a Newtonian gravitational field where the gateways are potential energy sinks. Within this approach, LAMs are selected, with a sample heuristic algorithm which performs a greedy selection based on the highest potential decrease. Note that this greedy approach has been chosen only for conceptual simplicity, but is not restrictive at all. Within a similar rationale, this heuristic strategy can be replaced with another technique of choice.

Instead of going for possibly suboptimal solutions of a simplified problem, we try to solve the original problem without any approximation. Moreover, our approach can be adapted to any interference model with only slight modification, but without changing the framework of potential energy and keeping a limited computational complexity. Simulation results derived with the ns2 simulator [9] confirm the goodness of the proposed strategy.

The rest of this paper is organized as follows. In Section II we formalize the problem and we describe the constraints of link activation modes. Section III introduces the original model of potential energy of a multi-hop network and relates it to routing and scheduling issues. A sample greedy algorithm is proposed which chooses the LAM which causes the highest potential energy decrease. This technique is numerically evaluated in Section IV. Finally, Section V concludes the paper.

II. PROBLEM STATEMENT

We consider a wireless mesh network where some nodes play the role of gateways, i.e., they are cabled to the Internet and can be seen as sink with very high capacity. We focus on algorithms to efficiently deliver a given amount of traffic over the network to the gateway nodes. In particular, we
concentrate on the minimal time scheduling problem, i.e., to deliver a given amount of traffic from all nodes to the gateways in the shortest possible time.

This is therefore an uplink case. The downlink case, i.e., the problem where the traffic is sent from the gateways to all nodes can be framed similarly, by considering reversed link directions and flipping the time axis. This problem is also closely related to the throughput maximization, i.e., to obtain the highest amount of traffic delivered to the gateways in an assigned time. Indeed, with minor modifications our framework can address this problem as well.

In the following, we will represent the multi-hop network as a graph \( G = (N, E) \). The wireless nodes are collected in set \( N \) and are connected by the edges belonging to set \( E \), thus representing the communication links of the network. The set \( V \subset N \) contains the gateways. We only consider connected graphs, where in particular a path exist from any node \( i \in N \) to at least one gateway \( j \in V \). However, differently from most related works we do not assume that the edges are necessarily bi-directional, i.e., the existence of \((i, j) \in E\) does not imply that also \((j, i) \) exists in the same set. This feature is only required when certain MAC protocols, such as IEEE 802.11 [10], are employed. This happens due to the exchange acknowledgement packets (ACK) at data link layer. When ACKs are sent, the logical receiver behaves also as a physical transmitter, and therefore links must be bidirectional.

To quantify the capacity of the link we make use of variables \( r_{ij} \), called link rates, which can be regarded as the amount of bits which can be transmitted over the link \((i,j)\) on a TDMA slot. We also consider a parameter \( g_{ij} \) corresponding to the wireless link gain over \((i,j)\). For each node \( i \in N \) we will refer to the backlog queue length at the node, assumed to be varying over time, as \( q_i(t) \). At time 0, all non-gateway nodes have a backlog of length \( q_i(0) \) to be sent to any of the gateways. The minimal time scheduling problem corresponds to finding the lowest length \( T_{\text{min}} \) of a feasible link activation pattern which delivers all traffic to the gateways. This means that

\[
T_{\text{min}} = \min \{ t : q_i(t) = 0, \forall i \in N \setminus V \}. \tag{1}
\]

For simplicity, we assume that the value of \( q_i(0) \) is known a priori and no further packet arrival takes place after link activation has started. In this way, if the uplink problem can be solved over a specified finite time-horizon \( T \), i.e. \( T_{\text{min}} \) is lower than or equal to \( T \), its solution can also serve as the basis for a periodic schedule, where a link activation pattern of length \( T \) is indefinitely repeated. A further extension is possible to the cases traffic with multiple priority classes or different required delay guarantees. Another option is to consider packet arrivals within the time frame. All these differences do not change most of the considerations we will present in the following, and can be investigated within a similar framework. We identify them as possible interesting directions for future research.

The problem of determining \( T_{\text{min}} \) exactly is very complicated. Not only the resulting optimization problem is NP-complete [2], but also it strongly depends on the network parameters, i.e., the graph topology, the edge rates and the initial backlog at each node. Solutions based on integer linear programming often introduce simplifications to make the problem more tractable, which we want to avoid.

Rather, following [7], we approach the JRS problem by determining a set of link activations. Link \((i,j)\) is said to be active if \( i \) transmits to \( j \). A group of links which are allowed to be simultaneously active form a LAM. It can be reasonably assumed that all transmissions belonging to the same LAM can be performed simultaneously in an error-free manner. To formally represent the LAMs within our graph-based approach, we define binary variables \( x_{ij}^{(m)} \) describing the activation of link \((i,j)\) in mode \( m \). We assume that \( x_{ij}^{(m)} \) is equal to 1 if LAM \( m \) includes the activation of link \((i,j)\), and 0 otherwise. The index \( m \) spans over a proper set \( \mathcal{M} \). With a slight abuse of notation, we will use the symbol \( m \) to indicate both a single feasible mode (i.e., a set of links) and its numerical index. Similarly, \( \mathcal{M} \) refers both to the list of all LAM and the set of all their indices.

The simultaneous activation of multiple links improves the transmission parallelism. To decrease the schedule length one should activate as many links as possible [3]. However, not all links can be activated in the same time slot.

There are two fundamental types of constraints that prevent links from being simultaneously activated. First of all, the radio equipment of a single node limits the number of simultaneous transmissions and receptions at the same node to one at most. Secondly, wireless interference may prevent some links between different nodes.

**Transceiver constraints** — The activation of links incoming at or exiting from the same node is limited by the physical capabilities of the transceiver. In this paper, we focus on narrowband channels, where it is not possible to receive simultaneously from multiple sources. We therefore assume that at most one signal can be decoded, thus there is no point in sending multiple transmissions to the same receiver. Indeed, the correctness of this reception is related to the impact of wireless interference, as will be discussed in the next subsection. Yet, regardless of the interference model, the maximum number of simultaneous successful receptions is one. A similar situation happens for the transmitter. Multicast transmissions, i.e., from one transmitter to many receivers, are actually possible on the wireless medium. However, the information content is the same for all receivers. For this reason, this situation is not relevant here. Multiple transmissions of different packets from the same node are instead forbidden. Finally, also transmissions and receptions at the same node can not happen in the same
time slot, since the transmitted power signal will destroy any packet reception [11]. In other words, the wireless communication medium is intrinsically half-duplex. Indeed, full-duplex capability could be obtained at the price of additional resource, e.g., by using directional antennas [12], which are however out of the scope of the present paper.

For these reasons, we impose that the activation of links should satisfy what we call half-duplex constraint, which corresponds to not activating more than one operation (i.e., either a transmission or a reception), for each node. Formally, this translates into:

\[ \forall i \in \mathcal{N}, \forall m \in \mathcal{M} \sum_{j \in \mathcal{S}_i} x_{ij}^{(m)} + \sum_{j \in \mathcal{R}_i} x_{ij}^{(m)} \leq 1, \quad (2) \]

where \( \mathcal{S}_i \) and \( \mathcal{R}_i \) are the set of the in-neighbors and out-neighbors of \( i \). Note that the protocol interference model [6] already includes this limitation. However, we emphasize that it is important to distinguish (2) from any kind of interference constraint, since it does not have to do with the wireless medium on which signals are transmitted, but with the limited capabilities of the terminal. The duplexing limitation holds irrespective of the interference model. For this reason, we will always impose the half-duplex constraint as a limitation to the parallelism of link activation which is independent of the radio interference.

Models for Interference Constraints — In [6], two useful models of interference among radio transmissions are introduced. Following this classification, we refer to them as protocol and physical interference model, respectively. Also, other extensions are available in the literature [13].

In this paper, we use the physical interference model, which is generally considered to be more realistic, but also more complex, than the protocol model. However, we emphasize this important aspect. The rationale of our analysis is not constrained to any peculiar aspect of a specific interference model. The only motivation of our choice is to show that our approach works in the most complicated case. On the other hand, any interference model can be used without changing the rationale, since it simply would end up in a different set of feasible LAMs. As the LAMs are determined a priori, this does not affect the selection strategy according to the potential energy framework that we will present in the following.

The physical interference model can be formalized in the context of LAM feasibility as follows:

\[ \sum_{k \in \mathcal{S}_i \setminus \{i\}} g_{kj} \sum_{\ell \in \mathcal{R}_k \setminus \{j\}} x_{ij}^{(m)} \geq \gamma_j \quad (4) \]

for any edge \((i,j)\) activated by mode \( m \), i.e., if \( x_{ij}^{(m)} = 1 \).

The key assumption of the model, i.e., the possibility to see the PER as a step function around a SIR threshold \( \gamma \), is indeed an approximation. Nevertheless, it is much more accurate than the ones made under the protocol models [2].

### III. Potential Energy

The goal of delivering a given amount of traffic to one or more gateways, has many similarities with the problems of water drainage which are present in civil engineering [8]. This is evident also from the terminology used, which often uses “sink” as a synonym for “gateway.” In the following, we will investigate the task of delivering a backlog \( q_i(0) \) (for brevity, in this section the time index will be often suppressed and we will speak of \( q_i \)) from any node \( i \) to one of the gateways within a potential energy framework, which imitates the representation of a Newtonian gravitational field.

In physics, the potential energy is a scalar function of the coordinates of an object, describing the energy that the object owns by virtue of the position within a force field. According to the nature of the force, it is associated with some physical properties of matter. For example, think of an object within a Newtonian gravitational field. In this case, potential energy is attributed to the object proportionally to its mass and height. Hereafter, we will always use this example as a reference case. We will therefore speak of a mass \( m \) located at height \( h \), which has a potential energy proportional to \( mh \). It is not restrictive to assume that the proportionality constant, which depends on the unit of measure, is equal to 1. Thus, the height is also the value of the scalar potential for the Newtonian gravitational field. The potential energy associated with a mass and a position also corresponds to the work to move the mass there from a position which is conventionally assumed to be located at zero height.
In our case, it is immediate to relate the mass with the amount of traffic which forms the backlog of a node. It is also reasonable to think of the gateways as potential sinks, i.e., positions at zero height. The underlying idea of our approach is that a mass (i.e., an amount of traffic) located at a given position (i.e., in queue at a given node) should be associated with a potential energy. We relate the potential energy \( \Pi \) to the delivery of \( q_i \) to one of the gateways. In particular, we define it as the minimum time to deliver \( q_i \) without any pipeline effect, i.e., if multiple hops are present, the field with the test mass itself. In this sense, “testing” within the force field. This must be done without perturbing the field with \( q \) means two things. First, any other traffic source must be turned off. In other words, \( q_i \) is the only traffic present in the network. Second, we assume that \( q_i \) is atomic, i.e., it can not be split over multiple links. This conditions does not necessarily mean that \( q_i \) contains a single packet: an atomic backlog can consist of multiple packets, but it can not be pipelined, i.e., as discussed above, it must be entirely received before being further retransmitted. If any of these conditions is violated, the evaluation of the potential energy will be no longer correct, as we must also take the compatibility of multiple transmissions into account.

If an atomic backlog \( q \) is sent through the series of two links having rates \( r_1 \) and \( r_2 \), the delivery time would be equal to \( q(i_1^{-1} + i_2^{-1}) \), i.e., the overall transmission rate is the harmonic average of the rates. This fact can be generalized to the series of any number of links (again, without pipeline effect). Hence, the path which would require the lowest time to transmit an atomic amount of traffic \( q_i \) from node \( i \) to a gateway, can be easily evaluated, e.g., by applying the well known Dijkstra algorithm taking the reciprocal of the rates as link weights. According to the reasoning above, these weights are non-negative and additive (i.e., they are summed over series of links). The path obtained in this way will be called in the following the \textit{fastest path to gateway of the atomic backlog} (FP2GAB) of node \( i \). The FP2GAB rate of node \( i \), i.e., the harmonic average rate evaluated over all links belonging to the FP2GAB, will be denoted as \( \rho_i \).

According to these reasoning, the potential energy \( \Pi_i \) of backlog \( q_i \) located at node \( i \) with FP2GAB rate \( \rho_i \) is equal to \( \Pi_i = q_i / \rho_i \). Adopting the same notation of the Newtonian gravitational field, it can also be written \( \Pi_i = q_i h_i \), where \( h_i \) is the height (scalar potential) of node \( i \). The height of node \( i \) results in this way equal to \( (\rho_i)^{-1} \). Correctly, \( h_i \) depends on position characteristics only, where “position” is meant in the topological sense.

### Table I

**Pseudo-code of the LAMPo-greedy algorithm**

An important observation is that when the network parallelism and/or pipeline effects are exploited, the minimal scheduling time can be lower than the overall potential energy of the network. However, both the scheduling time and the potential energy exhibit similar trends; in particular, they increase when the backlogs \( q_i \) are higher. Thus, we can determine a JRS solution through sequences of LAMs selected with respect to the impact they have on the potential energy of the whole network. We will describe a possible application of this approach, that we will call LAMPo (as a short for LAM Potential), in the next section.

Now we can address the evaluation of the number of slots required to transmit all the traffic to the gateways in a TDMA approach. We assume that the centralized network control determine a LAM to be performed for the entire duration of a single time slot. Remember that a feasible LAMs describes a set of links which can be activated together without violating half-duplex and interference constraints, thus we can reasonably assume that all involved transmissions successfully deliver their packets to the destination.

We need a criterion to decide which LAM to activate, and in which order. The reasoning behind the LAMPo approach is that when all packets have been delivered to the sinks, the overall potential energy of the network is equal to 0. On the other hand, the potential energy of the traffic at a node is by definition the delivery time on its FP2GAB. As discussed in the previous section, the potential energy of the entire network represents an upper bound on the delivery time for the whole network traffic. More transmissions may be activated in parallel to exploit the pipeline effect so as to decrease the schedule length [3].

```plaintext
function LAMPo-greedy
1. evaluate the FP2GAB for all nodes;
2. let \( \rho_i \) be the FP2GAB rate of node \( i \);
3. denote the set of all LAMs as \( \mathcal{M} \);
4. initialize \( t = 0 \) and the schedule \( \mathcal{L} = \emptyset \);
5. while \( \sum_{i \in \mathcal{N}} q_i(t) > 0 \)
6. evaluate \( \Pi = \sum_{i \in \mathcal{N}} q_i(t) \rho_i^{-1} \)
7. for mode \( m \in \mathcal{M} \)
8. \( q_i^* = q_i(t) \) for all \( i \in \mathcal{N} \); \#
9. for \( (i, j) \in \{ \text{all active links in mode } m \} \)
10. \( q_i^* = \max(0, q_i(t) - r_{ij}) \); \#
11. end-for over active links \( (i, j) \);
12. \# \# \Pi' = \sum_{i \in \mathcal{N}} q_i^* \rho_i^{-1} \#
13. if \( \Pi - \Pi' > \Delta \Pi \)
14. \# selected-mode = \# \Pi;
15. \# \# \Pi = \Pi - \Pi';
16. end-if;
17. end-for over modes;
18. update \( q_i(t) \) accordingly to selected-mode;
19. \#\ ADD SELECTED-MODE TO \mathcal{L} \#
20. return the schedule \( \mathcal{L} \);
```

An important observation is that when the network parallelism and/or pipeline effects are exploited, the minimal scheduling time can be lower than the overall potential energy of the network. However, both the scheduling time and the potential energy exhibit similar trends; in particular, they increase when the backlogs \( q_i \) are higher. Thus, we can determine a JRS solution through sequences of LAMs selected with respect to the impact they have on the potential energy of the whole network. We will describe a possible application of this approach, that we will call LAMPo (as a short for LAM Potential), in the next section.

Now we can address the evaluation of the number of slots required to transmit all the traffic to the gateways in a TDMA approach. We assume that the centralized network control determine a LAM to be performed for the entire duration of a single time slot. Remember that a feasible LAMs describes a set of links which can be activated together without violating half-duplex and interference constraints, thus we can reasonably assume that all involved transmissions successfully deliver their packets to the destination.

We need a criterion to decide which LAM to activate, and in which order. The reasoning behind the LAMPo approach is that when all packets have been delivered to the sinks, the overall potential energy of the network is equal to 0. On the other hand, the potential energy of the traffic at a node is by definition the delivery time on its FP2GAB. As discussed in the previous section, the potential energy of the entire network represents an upper bound on the delivery time for the whole network traffic. More transmissions may be activated in parallel to exploit the pipeline effect so as to decrease the schedule length [3].
There is necessarily a better mode that decreases the potential energy, e.g., by activating a single link which moves some traffic in the direction of the gateway across the FP2GAB. A good LAM to select for activation is one that decreases significantly the potential energy of the entire network. The higher the decrease achieved with a single LAM activation, the better the improvement to the transmission parallelism and therefore to the overall delivery.

Therefore, we derive a simple greedy strategy which selects the LAM achieving the highest decrease on the potential energy of the entire network. This strategy, that will be referred to in the following as LAMPo-greedy algorithm, can be described by the pseudo-code reported in Table I. At each iteration, the selected LAM to be added to the schedule \( \mathcal{L} \) is initialized as \( \text{mode}_0 \), since it is exploited that it causes a variation of \( \Pi \). Then, a greedy search is performed which updates this selected mode with the best one found over all LAMs.

IV. PERFORMANCE EVALUATION

In this section, we evaluate the performance of LAMPo framework in various relevant scenarios. First, we introduce the simulation environment and the performance indices analyzed. Then, we present the simulation scenarios and the results. The analysis was carried out by means of Network Simulator 2 (ns2) [9].

We consider a grid consisting of \( 30 \) m \( \times \) \( 30 \) m squares. Nodes occupy the grid intersections in a contiguous manner. We consider \( 2 \times 3, 3 \times 3, 3 \times 4 \) and \( 4 \times 4 \) grid dispositions of the nodes. We assume that there is only one gateway in the network (placed in a corner of the grid) and each of the other nodes has a fixed number of packets to transmit toward the gateway. The schedule is computed according to the LAMPo-greedy algorithm.

We perform 10 simulation run for any scenario. Each simulation run corresponds to a different instance of the network topology. In fact, even though the node placement is identical for any instance of the scenario, the channel has random behavior, for what concerns both gains \( g_{ij} \) and rates \( r_{ij} \). Thus, the obtained network topology is in general different for each simulation run.

Indeed, we assume that the channel gain of an edge having length equal to \( d \) consists of two terms, i.e., path loss and shadowing. While the former only depends on \( d \) and is therefore equal for the same link in any scenario instance, the latter has random behavior which depends on other factors than the distance. The path loss term is taken as proportional to \( d^{-3.5} \); without loss of generality, we can assume that the proportionality constant (i.e., the path loss at 1 meter) to be equal to 1. The shadowing term is a log-normal random variable with zero mean and standard deviation equal to 5 dB, however shadowing variables of different links are correlated through a two-dimensional extension of the Gudmundsons model [14], with a correlation factor equal to 0.6 at 100 meters. Note that usually the wireless channel gain is assumed also to have a fast fading component. This term, which is rapidly variable, may be taken into account as a fade margin in the SIR threshold \( \gamma \). The rate of a communication link \((i, j)\) is a discrete value function of the gain \( g_{ij} \). Table II reports the rate values assigned according to the attenuation with respect to the average path loss at 1 meter. If the gain falls within the range reported in the left-hand column of the table, the rate \( r_{ij} \) is equal to the value in the right-hand column, expressed in packets/slot.

![Figure 1. Schedule length with variable number of packets per node in several grid topologies.](image)

In this scenario we evaluate the impact of the network topology on the performance of the LAMPo-greedy algorithm. To this end, we vary the number of nodes in the network from 6 to 16 with different values of the nodes’ backlog. Here, we set the SIR threshold \( \gamma \) to the constant value of 2 dB, though the result is similar for other choices of \( \gamma \). In Fig. 1, we report the average schedule length versus the backlog per node in the case of different grid topologies. As can be seen, the average schedule length increases linearly with the number of backlogged packets per node for all the topologies considered. Furthermore, the greater the number of nodes in the network, the higher the value of the schedule length. This increase is roughly linear for low values of \( N \); it further increases when the network topology becomes larger and more bottlenecks can be present.

To further investigate the performance of LAMPo-greedy, we estimate the average end-to-end packet delay when the number of nodes in the grid ranges from 6 to 12. This is reported in Fig. 2. Each curve in the figures corresponds
to a specific number of backlogged packets per node. The average delay is shown to increase as the number of nodes in the network increases. Moreover, the greater the backlog of each node, the higher the packet delay as the schedule provided by the LAMPo-greedy algorithm requires a greater number of slots to deliver the overall network backlog to the gateway. The LAMPo-greedy strategy proves to scale well with respect to the amount of backlog per node. The curves increase almost proportionally to the node backlogs.

Finally, the average number of operations performed by the LAMPo-greedy algorithm is provided in Table III for the case of 300 backlogged packets per node and $\gamma = 2$. The number of operations grows exponentially, even though this is mainly due to the exponential increase in the number of the LAMs. In fact, compare the second and the third column of the table, which exhibit a similar exponential increase. Techniques to improve the efficiency of the LAM generation in order to obtain smaller (though non exhaustive) sets of feasible LAMs can be interesting goals of further research.

For what concerns the greedy selection algorithm itself, the complexity is indeed quite limited as the algorithm simply scan the list to find the largest decrease of the potential energy. This can be seen by considering the ratio, reported in the fourth column, between the actual number of operations and the list size. This value still increases in $N$ but in an approximately polynomial way (of the order of $O(N^2)$). Moreover, we also remark that the overall number of operations is in any case quite acceptable compared to current capabilities of microprocessors. In other words, these complexity values are highly competitive with respect to LP approaches which exhibit much higher complexity.

### V. CONCLUSIONS

In this paper, we analyzed the JRS problem in wireless mesh networks. We have proposed an approach based on LAMs, and introduced a novel framework called LAMPo, based on the definition of potential energy for multi-hop networks to solve the minimal time scheduling problem.

### REFERENCES


