
© 2009 IEEE.
Reprinted with permission.

In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of Aalto University's products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.
Approximating Maximum Directed Flow in a Large Wireless Network

Jarno Nousiainen  
Dept. of Communications and Networking  
Helsinki University of Technology  
P.O.Box 3000, FI-02015 TKK, Finland  
Email: jarno.nousiainen@tkk.fi

Pasi Lassila  
Dept. of Communications and Networking  
Helsinki University of Technology  
P.O.Box 3000, FI-02015 TKK, Finland  
Email: pasi.lassila@tkk.fi

Abstract—We study the maximum forwarding capacity for the relay traffic that can be transmitted through a wireless multihop network in a single direction. The problem appears as the microscopic level problem in a dense multihop network where the routing and forwarding tasks can be considered independently (separation of scales). Ultimately, the problem of finding the maximum forwarding capacity involves solving a maximum flow problem in an infinite plane with an infinite dimensional scheduling vector as an additional parameter to be optimized. In this paper, we approximate the infinite network by a finite but large network consisting of nodes distributed as a spatial Poisson process, and give the problem an LP formulation assuming a Boolean interference model. The computational complexity is further reduced by relaxing the necessary and sufficient constraints and solving the LP problem with a reduced set of necessary clique constraints. This gives a new significantly tighter upper bound on the achievable forwarding capacity compared with our previous (non-achievable) upper bound corresponding to the maximum capacity in one time slot.

I. INTRODUCTION

In the future applications of large wireless multihop networks, such as wireless sensor networks (WSNs), the size of the network may rise to several thousands of nodes. The exact modeling of large networks is problematic, and much of the available results are limited into scaling laws like the seminal work by Gupta and Kumar [7] and the recent results by Franceschetti et al. in [5]. However the scaling laws leave the constant of proportionality undetermined. We aim to provide more concrete results that fix the constant of proportionality.

We consider a wireless network in a given domain in the limit of a dense wireless network. In this limit, an elegant separation of scales takes place, and the problem of maximizing the capacity separates into two independent problems: a macroscopic and microscopic problem, see [9], [10]. At the macroscopic level a continuum approach can be adopted where routes are continuous geometric curves and the problem is that of load balancing [9], [10]. Similar continuum models have also been analyzed in [3], [11], [13]. At the microscopic level, corresponding to the scale of individual nodes, the network appears as an infinite random wireless network, and the task is to maximize the so-called forwarding capacity of the network. While practical wireless networks cannot be considered infinitely dense, the results obtained at this limit offer useful information on efficient routing and coordination schemes in finite networks as well.

The focus here is on the microscopic level problem. The network consists of nodes distributed according to a spatial Poisson process, and interference is modeled using the so-called Boolean interference model. We operate in a parameter area where most of the nodes are connected and form what we call a transport network. Full connectivity is not required at this level of operation. Assuming that the traffic flowing in different directions can be handled via time sharing, the specific task is to maximize the forwarding capacity of the network in a given direction by optimal coordination of transmissions. The forwarding capacity of the system is expressed as the mean density of progress, see [1], [16].

The problem has a similar fundamental nature as finding the percolation threshold, which has been studied for wireless networks recently under various interference models in [4]. The forwarding capacity remains zero until the network percolates, but instead of the point where the network becomes connected, we are interested in the point where the forwarding capacity is maximized.

In our previous paper [15], certain upper bounds on the forwarding capacity were already established, of which the tightest represents the maximum achievable density of progress in a single time slot. This (non-achievable) upper bound is relatively loose, because it gives a positive value as soon as links begin to emerge and before there is any kind of long range connectivity in the network.

In this paper, the infinite network is approximated by a large finite network where the capacity problem can be formulated as a maximum flow problem. The flow problem can be represented as an LP problem by using two different approaches, i.e., by using the so called transmission mode or the clique formulation. Both approaches are computationally hard to solve due to their inherent NP completeness. While the transmission mode formulation characterizes the necessary and sufficient conditions for flow optimality, the clique formulation only yields necessary conditions. Both approaches have been used widely to analyze the capacity of finite wireless multihop networks under different assumptions regarding, for example interference models, routing and channel assignment, see, e.g., [12], [14]. Typically, the problems are solved using heuris-
tics. Also, so called FPTAS algorithms have been developed that guarantee a fully polynomial time complexity and that the difference between the obtained solution and the actual optimum is bounded by a given factor, see [14]. In [8], algorithms applying so called scaled clique constraints have been developed to yield necessary and sufficient conditions.

We specifically apply the LP formulation using the clique constraints to address the forwarding capacity. This results in a certain formulation for the maximal flow problem where the sources and sinks lie at the opposite edges of a large but finite square-shaped network, and traffic is relayed through the network. A heuristic polynomial time algorithm is derived for approximating the complete set of constraints resulting from identifying all maximal cliques. The approximation still yields an upper bound for the total achievable forwarding capacity. This allows us to compare the single slot (non-achievable) upper bound, as derived in our earlier paper [15], to the new upper bound on the achievable forwarding capacity. The new results significantly tighten the previous results on the upper bounds of the forwarding capacity.

The remainder of this paper is structured as follows. In Section II, we define the problem decomposition and in more detail the microscopic level problem. Section III gathers together the graph and LP formulations and presents the heuristics used in determining the constraints. The numerical results are in Section IV, while Section V concludes the paper.

II. PRELIMINARIES AND ASSUMPTIONS

This section ties the problem of finding the maximum forwarding capacity of an infinite wireless multi-hop network into a broader problem of maximizing the total data flow in dense multihop networks.

A. Problem decomposition

In a large wireless multihop network, a path between a sender and a receiver consists of a large number of hops. In the limit of a dense network [9], [10], maximizing the capacity of the network separates into two problems, as detailed below.

A dense network corresponds to a network in a closed domain with nodes having an infinitesimal transmission range, and the paths being smooth geometric curves allowing a continuous representation of the network. The traffic demand density profile \( t(r_1, r_2) \) \([1/m^2]\) is the fraction of the total rate of packet flow \( \Lambda \) in the whole network that originates from a differential area element about \( r_1 \) and is destined to a differential area element about \( r_2 \). With a given traffic profile and a set of paths \( \mathcal{P} \), the traffic load at \( r \) is denoted by \( \Psi(r, \mathcal{P}) \).

Forwarding capacity \( I \) sets the limit

\[
\Lambda \Psi(r, \mathcal{P}) \leq I \quad \forall r \quad \Rightarrow \quad \Lambda \leq \frac{I}{\max_{r} \Psi(r, \mathcal{P})} \tag{1}
\]

for the local traffic load, and in order to maximize the capacity of the network, i.e., solve \( \max \Lambda \), we have two separate problems: 1) maximize the forwarding capacity, \( I^* = \max I \), and 2) minimize the maximal load.

Minimizing the maximal load (problem 2 above) corresponds to load balancing and is referred to as the macroscopic problem. Determining the maximal forwarding capacity \( I^* \) (problem 1 above) represents the microscopic level problem capturing the properties of the underlying wireless network from the point of view of a single node.

B. Microscopic level problem

In this paper we focus on the problem of determining \( I^* \), while the load balancing problem is discussed, e.g., in [9], [10]. At the microscopic level, the assumption of a dense network implies that from the local perspective the network appears as an infinite network. Two randomly selected nodes are, on the average, much further apart from each other than two neighboring ones. If the nodes communicating with each other are assumed to be random, a route between a source and a destination typically consists of a large number of hops. Therefore, the amount of relay traffic in a specific area of the network is much higher than the amount of traffic that originates from or terminates to the area.

When considering the relay traffic, we do not require full connectivity, but the nodes that do not belong to the transport network can be connected to it via an access network in negligible time, since the amount of originating and terminating traffic in an infinite network is very small compared to the amount of relay traffic, cf. [5]. This allows us to concentrate purely on the relay traffic and omit the originating and terminating traffic from the model. No traffic matrix or distribution is needed for determining \( I^* \), but we simply maximize the amount of traffic that can be relayed through the network.

In a network there is traffic flowing in different directions. However, assuming that time is slotted, the different directions can be treated independently so that the maximal forwarding capacity \( I^* \) can be shared between the different directions using appropriate scheduling.\(^1\) Thus, the microscopic level problem considers maximizing the flow of traffic in a given direction.

C. Forwarding capacity

The forwarding capacity \( I^* \) is defined as the maximum sustainable density of packet flow \([1/m/s]\), i.e., the number of packets crossing a unit length of a line perpendicular to the flow per unit time. Alternatively, \( I^* \) can be interpreted to represent the maximum sustainable mean density of progress [16], i.e., the density of packets times their mean velocity in a given direction. The maximum sustainable density of flow (obtained with optimal global coordination of the transmissions) depends on the physical parameters at hand: density of nodes \( \lambda \) \([1/m^2]\), transmission range \( R \) \([m]\), and nominal capacity of a link \( C \) \([1/s]\). By dimensional analysis, \( I^* \) can be expressed as any combination of the parameters having the dimension \( 1/m/s \) times a function of all the independent dimensionless parameters that can be formed. A combination of parameters

\(^1\)The direction distribution in different areas of the network might be different. This mismatch becomes negligible when the size of the network approaches infinity.
of dimension $1/m/s$ is provided by $C\sqrt{\lambda}$, and there is only one dimensionless parameter, namely the mean degree of a node $\nu = \pi R^2$ (the constant $\pi$ is unimportant as it can be absorbed in the definition of $u$ below). Thus,

$$I^* = C\sqrt{\lambda} u(\nu),$$

where $u$ is an unknown function referred to as the dimensionless mean progress. Our task is to find this $u(\nu)$, or to be more specific, an upper bound for it. Additionally, there is a given value $\nu^*$ (or, equivalently, for a given $\lambda$ the optimal transmission range $R^*)$ that maximizes the function $u$.

Note that the famous scaling law $O(1/\sqrt{n})$ of the network capacity per node [7] follows trivially from this relation together with (1), noting that for a given domain and a given traffic demand profile $t(r_1, r_2)$ the denominator of (1) is a given constant. However, the validity of (1) is based on the physically intuitive notion of separation of scales at high node densities, for which we have not given a rigorous proof.

III. LP FORMULATIONS

In this section, we define the notations required to formulate the problem as a linear programming problem as well as the actual LP formulations and the heuristics used to reduce the computational burden.

The first LP formulation is the complete schedulable max-flow problem for the transmission modes, while in the second one these sufficient conditions have been replaced with constraints for the clique capacities, or to be more exact the flow problem for the transmission modes, while in the second set of constraints, it is possible to obtain results from finite network realizations. Finally, an example illustrates the compromises that have to be done in order to produce an upper bound for the forwarding capacity of the network.

A. Max-flow problem

When the nodes correspond to the vertices, $V$, of a directed graph $G = (V, E)$, there exists an edge $(u, v) \in E$, $u, v \in V$ if $d(u, v) \leq R$, where $d(u, v)$ is the Euclidean distance between the nodes. In the Interference Graph, $IG$, the links of the network correspond to vertices, and two vertices are adjacent if the corresponding links interfere with each other, that is, links $a$ and $e$ are adjacent if

$$d(t(a), r(e)) \leq R \lor d(r(a), t(e)) \leq R,$$

where $t(e)$ is the transmitting node of link $e \in E$ and $r(e)$ the receiving node. This interference model is referred to as the Boolean interference model which says that a node is only capable to receive a transmission if it is the only one within its transmission range.

With the term network we refer to a pair $(G, c)$ of a graph and a mapping $c : E \rightarrow \mathbb{R}^+$. In general, $c(e)$ is called the capacity of the edge $e$. In a wireless network, not all links can be active simultaneously due to interference, and thus, in our model for a given link $e$, the effective link capacity $c(e)$ is less than the nominal capacity $C$. To define $c(e)$, we have to establish a schedule $\alpha$ which tells us how the links are used. All the links that are active simultaneously have to belong to the same independent set of links to avoid collisions. A set of links $L$ is said to be independent if the corresponding vertices form an independent set in $IG$. This is equivalent with forming a clique in the complement of $IG$. A complement of a graph has the same vertices as the original graph, and two vertices are adjacent iff they are not adjacent in the original graph. The independent sets of links (transmission modes) are cliques in the complement interference graph. A clique is a set of vertices such that for every pair of vertices there is an edge connecting them.

We call the independent sets that are used for transmitting modes and denote the set of transmission modes with $M = \{\mathcal{L}_1, \ldots, \mathcal{L}_M\}$. The schedule $\alpha = \{t_1, \ldots, t_M\}$ assigns each transmission mode $\mathcal{L}_i$ with the proportion of time $t_i$ that it is used. Now the effective capacity of link $e$ is

$$c(e) = C \sum_{i=1}^{M} t_i \mathbf{1}_{\{e \in \mathcal{L}_i\}},$$

that is, the nominal capacity times the time share the link is active.

Let us now consider a network $(G, c)$ where we distinguish two sets of special vertices: the start nodes $S$ and the terminal nodes $T$, or the sources and the sinks. In our construction the sources are located in the left side of the transport network and the sinks on the right. Now, we have a structure $\mathcal{N} = (G, c, S, T)$ that we call a flow network, and we can define a flow in the network. A mapping $f : E \rightarrow \mathbb{R}_+$ is a flow if it satisfies the following conditions:

$$f(e) \in [0, c(e)) \quad \forall e \in E$$

$$\sum_{r(e) = u} f(e) = \sum_{t(e) = v} f(e) \quad \forall u \in V \setminus (S \cup T)$$

The first, feasibility condition guarantees that there is a positive ($\geq 0$) bounded flow through every arc, and the second, flow conservation condition means that flows are preserved (except at the sources and the sinks). The value of flow $f$ is

$$w(f) = \sum_{t(e) \in S} f(e) - \sum_{r(e) \in S} f(e) = \sum_{r(e) \in T} f(e) - \sum_{t(e) \in T} f(e).$$

The maximum flow problem is a classic problem in graph theory and combinatorial optimization with a variety of applications. It considers finding a feasible flow through a flow network that is maximal. A flow $f$ is maximal if $w(f) \geq w(f')$ for all flows $f'$ on $\mathcal{N}$.  

B. Transmission mode formulation

The problem is to maximize the value of the flow, $w(f)$, subject to the flow constraints (5-6) and a constraint that says that the sum of the time shares allocated to the transmission
modes cannot exceed one. Hence, the decision variables are the flow, $f$, and the schedule, $\alpha$, that specifies the capacities.

$$\max_{f, \alpha} w(f) \quad \text{s.t.} \quad \begin{cases} 0 \leq f(e) \leq C \sum_{i \in E} t_i 1_{\{e \in C_i\}} & \forall e \in E \\ \sum_{r(e)=v} f(e) - \sum_{r(e)=v} f(e) = 0 & \forall v \in V \setminus (S \cup T) \\ \sum_{i=1}^M t_i \leq 1, t_i \geq 0, & i = 1, \ldots, M \end{cases}$$

The previous constraints are the necessary and sufficient conditions for the flow optimality.

C. Clique formulation

Each link in the network corresponds to a vertex in the interference graph, and two vertices are adjacent if the corresponding links interfere with each other. From the constraint that says that the sum of the time shares of the transmission modes is less than or equal to one, we get that the total capacity of the links forming a clique in the interference graph never exceeds the nominal capacity $C$. This follows from the fact that adjacent vertices (interfering links) belong to different transmission modes, and thus cannot be activated at the same time. Vertices in a clique are all pairwise adjacent, and hence only one of them can be used at a time.

This means that a constraint that limits the capacity of a maximal clique in the interference graph to $C$ is a necessary condition for a feasible flow.\(^2\) By replacing the constraint considering the time shares with constraints for the clique capacities, we get an upper bound for the maximum value of a flow in the network. Since the flow through a link is bounded by the capacity, these constraints can further be replaced by similar ones considering the flow. When the flow satisfies these, the capacities could always be chosen to match the flow. Thus, the only decision variable is the flow.

$$\max_f w(f) \quad \text{s.t.} \quad \begin{cases} f(e) \geq 0 & \forall e \in E \\ \sum_{r(e)=v} f(e) - \sum_{r(e)=v} f(e) = 0 & \forall v \in V \setminus (S \cup T) \\ \sum_{e \in C_i} f(e) \leq C & \forall \text{cliques } C_i \text{ of the IG} \end{cases}$$

When the diameter of the network grows compared to the transmission (interference) range, the number of transmission modes becomes very large while the growth in the number of cliques is much modest. In practice this means that this latter formulation becomes more tempting when the network is large, though it only gives an upper bound.

D. Clique Approximation Algorithm

The problem of finding a maximum clique is NP-hard [6]. It is computationally equivalent to finding a maximum independent set through the concept of complement graph. This means that though the latter formulation for the LP problem is simpler, it is still not usually possible to generate the constraints explicitly for a given large network in a feasible time.

\(^2\)Clique constraints are necessary and sufficient for perfect graphs [8].
might be missed. Not share a common clique, some of the pairwise dependencies to describe the interference caused by a single clique in the number of interfering paths grows, the algorithm is only able as the number of interfering paths is only two. When the to produce the same result with the full clique formulation clearly visible. The clique approximation algorithm is still able seen in the example where the upper and lower paths are to another to begin to interfere with each other. This can be "paths" used to transfer packets from one side of the network to differ when the mean node degree is large enough for the modes and for the clique formulation of the problem start 74 maximal transmission modes.

The number of maximal cliques is 27, and there exists a total of 74 maximal transmission modes.

The results for the original problem with transmission modes and for the clique formulation of the problem start to differ when the mean node degree is large enough for the “paths” used to transfer packets from one side of the network to another to begin to interfere with each other. This can be seen in the example where the upper and lower paths are clearly visible. The clique approximation algorithm is still able to produce the same result with the full clique formulation as the number of interfering paths is only two. When the number of interfering paths grows, the algorithm is only able to describe the interference caused by a single clique in the interference graph per forward link. If the interfering paths do not share a common clique, some of the pairwise dependencies might be missed.

### IV. Results

At the microscopic level the network appears as an infinite one. The idea is approximate this network with a finite, but a large, one where the relay traffic flows, e.g., from left to right. A unit square with the average of \( N \) nodes distributed according to a spatial Poisson process represented the relay network, and a strip of sources was laid on the left side of the square and a similar strip of sinks on the right. One transmission range \( R \) is a sufficient width for the strips since no longer links can be established. Finally, the top and the bottom of the network were connected to reduce harmful border effects. An example network with \( \nu = 10 \) and \( N = 1000 \) is depicted in Fig. 3. After solving the optimization problem (9) with the reduced set of clique constraints from the algorithm of Table I, we get an upper bound for the maximum flow from left to right. Since the network resides in a unit square, we have \( u = \frac{W(f)}{\sqrt{N}} \).

To properly simulate the operation of an infinite network, \( N \) needs to be as large as possible. Due to the computational limitations, it is not possible to solve the problem for sufficiently large networks when the mean degree of a node, \( \nu \), is large. Because it is easier to establish a flow through shorter paths, the price to pay for a smaller network is that we get a looser upper bound. The border effect caused by the left and the right sides of the network, where the sources and sinks are located, is illustrated in Fig. 4 which shows the obtained dimensionless mean progress \( u \) as a function of the system size \( N \) for \( \nu = 7 \) and 10. The figure also shows the rate at which the network starts to resemble an infinite network.

Fig. 5 represents \( u \) as a function of \( \nu \) obtained using the clique approximation algorithm in Table I with \( N = 500 \) as well as the 90% confidence intervals. The figure also depicts the maximum capacity achievable in one time slot [15] and a result from an actual forwarding method, namely the opportunistic forwarding method [15], that is a variant of ExOR [2]. As can be seen from the figure, the upper bound provided by the LP approximation is much tighter than the one-slot maximum when the mean degree of a node is small. On the other hand, when \( \nu \) increases even the relaxed LP problem becomes too complex, and the earlier results remain as the only available bounds.

We finally comment on the relationship between the one-slot maximum and the achievable flow for large \( \nu \). The maximum total progress in one time slot gives an upper bound for the total progress achievable with feasible flows. This is intuitive since in a continuous flow the best possible progress cannot be
achieved in every time slot. When the mean degree of a node is relatively small, the upper bound should be loose. For example, just before the network becomes connected the feasible flow is still zero, but there are plenty of links that can contribute to the one slot maximum that is quite high. When the mean number of neighbors increases, the one-slot approximation becomes better since the difference to the bottleneck-slot limiting the feasible flows decreases. This is due to the idea that when the network is very dense, it is always possible to use a relatively good transmission mode.

V. CONCLUSIONS

The potential of large wireless multihop networks as a part of the future ubiquitous networks is undeniable, but the relevant work is still heavily biased towards certain types of results, e.g., scaling laws leaving the constant of proportionality undetermined. To get concrete quantitative results that fix the unknown constant, one often has to rely on approximations due to the computational complexity of the problems.

We studied the directed max-flow problem in an infinite network that appears as the microscopic level problem when the separation of scales in a dense wireless multihop network is assumed. The other part of the problem needed to fix the constant, so-called load balancing problem, has been studied elsewhere. By solving the corresponding problem in a finite network with small mean degrees of a node, we were able to considerably reduce the gap between the earlier upper bound, i.e., the one-slot maximum capacity, and a capacity achieved with an actual forwarding method. In our case, the necessity of approximations comes from the NP hardness of generating the constraints for the LP problem. Hence, the original schedulable link capacity constraints of the max-flow LP problem were replaced with a reduced set of clique constraints generated with a heuristic algorithm.

The maximum achievable progress in one time slot still remains as the tightest upper bound with larger neighborhoods. The results also give some kind of indication of when the finite network starts to behave as an infinite one.

ACKNOWLEDGEMENTS

This research was conducted under the ABI project financially supported by TEKES. We are grateful for Prof. Jorma Virtamo for valuable comments.

REFERENCES