On the Macroscopic Effects of Local Interactions in Multi-hop Wireless Networks

V. Saligrama and D. Starobinski

Abstract— The objective of the paper is to provide qualitative insight into the global effects of distributed mechanisms, such as carrier sense multiple access (CSMA) and rate control, on the performance and stability of multi-hop wireless networks. Toward this end, we introduce a linear queueing network model where the service capacity of each node is modulated by the transmission state of its neighbor. We derive lower bounds on the steady-state utilization at each queue of such networks and demonstrate the existence of a phase transition phenomenon, whereby infinitesimal traffic increase at a single node in the network can suddenly render the entire network instable. We also present NS simulation results that show how this phenomenon can actually take place in IEEE 802.11 multi-hop wireless networks. Our results have direct bearing on rate control schemes, in that they indicate a minimum admissible threshold rate required to prevent network instability.

Fig. 1. Influence propagation in a 6-node network.

I. INTRODUCTION

One of the key reasons for the success of IEEE 802.11 (Wi-Fi) networks is their reliance on simple, distributed mechanisms. For instance, channel access is implemented using a variant of the carrier sense multiple access (CSMA) protocol that allows a user to transmit if it senses the channel to be idle [2]. Similarly, IEEE 802.11 networks support rate control mechanisms, whereby users can adapt their transmission rate based on local channel conditions (see [4] and references therein).

While the impact of the above mechanisms in single-hop (star) networks is pretty well understood (see, e.g., [3]), they can lead to much more complex interactions in multi-hop (mesh) networks. For example, consider the 6-node network shown in Fig. 1 and assume that each node implements CSMA. In this topology, node C cannot hear node A's transmissions, i.e., node A is a hidden node with respect to node C. Thus, if node A transmits whenever node C transmits, a collision will occur at node D [1]. In this scenario, node C will have to retransmit until a packet is successfully received at node D. These retransmissions by node C will in turn limit the fraction of time node E can successfully transmit to node F, thus triggering an even larger number of retransmissions by node E, and so forth. This example can be generalized to the case where nodes perform rate adaption, e.g., a node can
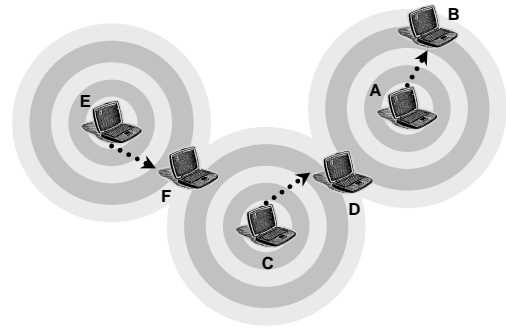
[1]The IEEE 802.11 standard includes an optional mechanism, called *RTS/CTS*, that can be used to reduce collisions due to hidden nodes, but is often not implemented in practice as it incurs significant overhead and does not fully prevent packet collisions [8].

receive data at a high rate μ when no nearby node interferes and at lower rate $k\mu$, where $0 \leq k < 1$, when a nearby node interferes (note that $k = 0$ corresponds to the hidden node case). In this case, because of a propagation effect similar to the one described above, it becomes increasingly likely that nodes down the chain can only communicate at the low data rate.

The objective of this paper is to get further insight into these propagation effects and better understand the impact of distributed channel access and rate adaptation mechanisms on the performance and stability of multi-hop wireless networks. Because of the difficulty of analyzing general multi-hop wireless networks, we focus our attention to the case where nodes are arranged along a linear topology. This configuration will occur repeatedly at different times and locations in large multi-hop networks.

In order to characterize the interaction between different nodes, we introduce a model called influence queueing network (IQN). Using this model, we identify situations under which traffic increase at a single node in the network can render the entire network instable. Moreover, we show that a distant node can potentially experience a *phase transition*, where virtually no change in the node's utilization is observed until it is suddenly forced to instability. We illustrate the existence of this phase transition phenomenon with several numerical examples. Furthermore, we present NS simulation results that show how the same phenomenon can actually take place in an IEEE 802.11-based wireless ad-hoc network.

Our analysis also provides key insight into the situations leading to network instability. Specifically, it reveals that instability is more likely to take place when the range of available transmission rates is large. Thus, to avoid instability, a decentralized rate control scheme should not allocate rates below a minimum admissible threshold level.

In Section II, we introduce and justify our IQN model and discuss related work. In Section III, we derive lower bounds on the steady-state utilization at each queue of these networks and in Section IV we demonstrate the existence of a phase transition phenomenon. These findings are validated by numerical results that are presented in Section V. In the same section, we also present simulation results obtained using the Network Simulator (NS) [1], showing phase transition in a multi-hop IEEE 802.11b wireless ad-hoc network. We conclude the paper in Section VI.

II. MODELS AND NOTATIONS

In this section, we first present our model, based on influence queuing networks, that provides an analytical framework to capture interaction between neighboring nodes in a multi-hop wireless network. We then provide justifications to our modeling assumptions and discuss related work.

A. The Influence Queue

We model the behavior of each node using an *influence queue*. According to this model, data packets arrive according to an independent Poisson process with mean rate λ and packets length are i.i.d. following a general distribution with mean $1/\mu$. Upon arrival, packets are either served immediately or forced to wait in an infinite-length buffer (if the server is found to be busy). Packets in the buffer are served according to a "first in, first out" (FIFO) priority discipline.

At first glance, the influence queue model appears quite similar to the familiar $M/G/1$ queuing model. However, an influence queue includes two additional components: an influence process, $X(t)$ and a capacity process $C(t) = f(X(t))$. The influence process at each node is a function of the states of other nodes in the network.

We now focus on the most basic example of an influence queue model, the binary-rate influence queue. First, let the influence process $X(t) \in \{0, 1\}$ be an on-off process. The capacity process (or the modulated process) is defined as

$$C(t) = \begin{cases} 1 & X(t) = 0 \\ k & X(t) = 1, \end{cases} \tag{1}$$

where $0 \leq k < 1$ is the influence rate. In other words, the queue operates at rate k when the influence process is "on" and at unity when it is "off". In this paper, we apply this model to the case where the transmission rate of a node depends on the transmission state of a neighboring node. Specifically, if the neighbor transmits at time t, then $X(t) = 1$; otherwise, $X(t) = 0$.

B. Linear Influence Queuing Networks

An influence queuing network (IQN) is a network in which each node is represented by a single influence queue, and the influence between these nodes is defined by the state of one or more other queues in the network.

We next focus on a special class of IQNs, referred to as *linear* IQN, that is defined as follows. Consider an infinite network of influence queues with arrival rates λ_n and packet
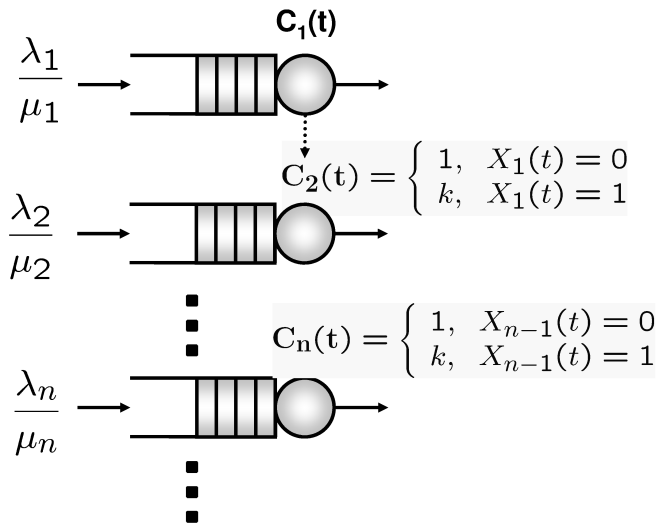

Fig. 2. A linear influence queuing network.

length $1/\mu_n$, $n = 1, 2, \dots$. The steady-state utilization of the nth queue is denoted by ρ_n. Queue 1 is "un-influenced", i.e. its capacity process $C_1(t) = 1$ for all t. Each of the remaining queues is directly influenced only by the preceding one, and influences only the following one. Thus, the service capacity process of queue n for $n = 2, 3, \dots$ is:

$$C_n(t) = \begin{cases} 1, & X_{n-1}(t) = 0 \\ k, & X_{n-1}(t) = 1, \end{cases} \tag{2}$$

where $X_n(t)$ is an on-off process that depends on whether or not the preceding $n-1$th queue is transmitting. An illustration of a linear influence queuing network is shown in Fig. 2.

Our primary questions of interest for linear IQNs are as follows. First, if we were to increase the mean traffic rate of only the first queue, how would this influence propagate to affect the new steady-state utilization of the other queues in the network? Second, under what conditions are we guaranteed that simply changing the traffic load on the first queue will drive some queue in the network to instability, especially for very distant queues ($n \rightarrow \infty$)? Finally, how does this behavior change as a function of the influence parameter k and the initial load on the network?

C. Model Justification and Related Work

The Poisson process, in addition to making the analysis tractable, can be a suitable model for applications where traffic is generated by the aggregation of a large number of independent streams.

Our model assumes that interferences are only due to transmissions from an immediate upstream neighbor. This is a reasonable assumption in situations similar to the one depicted in Fig. 1 where the transmission by a nearby node dominates the interferences at the receiver. For instance, in that figure, node A is the dominant interferer for node D and node C is the dominant interferer for node F, etc. If the path loss

exponent α is high enough, e.g., $\alpha = 4$, then the interference caused by the immediate upstream neighbor will typically be much larger than the aggregate interference caused by all the other transmitting nodes.

Our model is closely related to priority queues [5] but differs in some key aspects. First, a priority queueing system has a single server and only a single queue can be served at any time. In contrast, in our model several queues can be served simultaneously at different rates. Second, unlike a priority queueing system, where the highest priority queue blocks all the other queues, in our model each queue is influenced only by one adjacent queue.

While problems involving influence queueing networks have been studied in the literature in the past, previous work has mostly focused on determining the stability region and devising scheduling policies that stabilize the network whenever the arrival rates are within this region [7, 10]. In contrast, our paper investigates how distributed channel access and rate control mechanisms can affect the stability of IQNs.

III. A General Lower Bound on the Server Utilization in Linear IQNs

The derivation of a closed-form expression for the steady-state utilization ρ_n of each queue in a linear IQN appears to be a difficult problem. Instead, we provide a general lower bound $\hat{\rho}_n$ on this quantity. This lower bound will be useful to determine the possible situations in which phase transition phenomena may take place.

Theorem 1 *Set $\hat{\rho}_1 = \rho_1 = \min\{\lambda_1/\mu_1, 1\}$ and for all $n \geq 2$*

$$\hat{\rho}_n = \min\left\{ \frac{\lambda_n}{((1 - \hat{\rho}_{n-1}) + k\hat{\rho}_{n-1})\mu_n}, 1 \right\}. \quad (3)$$

Then $\hat{\rho}_n$ is a lower bound on the steady-state utilization ρ_n at each queue n.

Proof: The proof is by induction. For $n = 1$, the claim is obviously true. Our goal now is to show that for any packet j in queue $n \geq 2$, we have the following lower bound on the expected service time

$$E[S_n(j)] \geq \left\{ \frac{1}{((1 - \hat{\rho}_{n-1}) + k\hat{\rho}_{n-1})\mu_n}, 1 \right\} \triangleq \frac{1}{\hat{\mu}_n}, \quad (4)$$

where $S_n(j)$ represents the service time of packet j at queue n. Since the above lower bound does not depend on j, it applies also to the steady-state service time S_n. Thus, once we prove Eq. (4), the proof of the theorem follows immediately since $\rho_n = \min(\lambda_n E[S_n], 1)$.

Now, suppose that for any packet j in queue $n - 1$ we have a lower bound on the expected service time independent of packet j, i.e.,

$$E[S_{n-1}(j)] \geq \frac{1}{\hat{\mu}_{n-1}}.$$

We are left to establish, based on the induction hypothesis, that a lower bound independent of any packet can also be established for queue n.

To simplify the exposition, we first consider the case of $k = 0$, which can be considered as the limiting case of a queue transmitting at an infinitesimally small data rate when influenced by another queue.

We point out two subtle aspects of our analysis:
1) Packets in queue n depart only when queue $n - 1$ is idle.
2) A packet in queue n enters service as soon as no other packets are ahead of it in the queue.

Therefore packets in queue n are being served (albeit at zero rate) even when queue $n-1$ is busy. Figure 3 illustrates such a situation where packet C in the second queue enters service at time instant v. We note that even though the packet gets served at non-zero rate only starting from time w, it does reduce the transmission rate of the server of the third queue starting from time v.

We now introduce some notations to keep track of various variables. We denote by $T_n(j)$ the *system time* spent by packet j at node n, i.e., the total time spent by packet j in the queue and server of node n. We denote by $W_n(j)$ the *waiting time* for packet j at node n, i.e., the time spent by packet j in the queue of node n. The *service time* is the difference between the system time and waiting time, i.e., $S_n(j) = T_n(j) - W_n(j)$. In Figure 3 the time interval $[v, z]$ corresponds to the system and service times for packet C (observe that, in this case, the waiting time is zero). Let $R_n(j)$ denote the *residual time* it takes to serve all the packets that are already in the queue of node n when packet j arrives. Let $N_n(j), M_n(j)$ be the number of new packets that arrive in queue $n - 1$ during the times $T_n(j), W_n(j)$ respectively. For packet C in Figure 3, $N_n(C) = 1$ and $M_n(C) = 0$.

The *workload* for packet j is the time packet j is served at the full rate (this corresponds to the intervals $[w, x]$ and $[y, z]$ for packet C). Finally, let $V_n(j)$ denote the sum of the workload and the initial *lead time* after which packet j in queue n is served at the full rate. This initial lead time accounts for the situation where packet j arrives at a time instant when no other packets are ahead of it in queue n, but there are packets in queue $n - 1$ that are to be served (this corresponds to the interval $[v, w]$ in Figure 3 for packet C). Observe that if there are packets in the server of queue n when packet j arrives, then the initial lead time is equal to zero for packet j.

Based on the above definitions we have,

$$T_n(j) = R_n(j) + V_n(j) + \sum_{m=1}^{N_n(j)} S_{n-1}(m),$$

$$W_n(j) = R_n(j) + \sum_{m=1}^{M_n(j)} S_{n-1}(m).$$

Therefore,

$$S_n(j) = V_n(j) + \sum_{m=1}^{N_n(j)} S_{n-1}(m) - \sum_{m=1}^{M_n(j)} S_{n-1}(m). \quad (5)$$

We now note that the expected value of $V_n(j)$ is no smaller than $1/\mu_n$. This is because $V_n(j)$ includes the lead time as well as the workload, which is equal to $1/\mu_n$. Therefore,

$$E[V_n(j)] \geq 1/\mu_n.$$

Taking expectations in Equation (5) we have,

$$E[S_n(j)] \geq \frac{1}{\mu_n} + E\left[\sum_{m=M_n(j)+1}^{N_n(j)} E[S_{n-1}(m)|N_n(j), M_n(j)]\right],$$

where we have used the property of nested expectations. Now, we remind that the arrival process into queue $n-1$ is a Poisson process that is independent of the service time of packets. Therefore $E[S_{n-1}(m)|N_n(j), M_n(j)] = E[S_{n-1}(m)]$. Putting these facts together and using the induction hypothesis $E[S_{n-1}(m)] \geq 1/\hat{\mu}_{n-1}$, we obtain

$$E(S_n(j)) \geq \frac{1}{\mu_n} + \frac{E(N_n(j) - M_n(j))}{\hat{\mu}_{n-1}}. \quad (6)$$

From the fact that the arrivals form a Poisson process, we have

$$E[N_n(j)] = E[E[N_n(j)|T_n(j)]]$$
$$= E[\lambda_{n-1}T_n(j)] = \lambda_{n-1}E[T_n(j)].$$

Similarly, we have

$$E[M_n(j)] = \lambda_{n-1}E[W_n(j)].$$

Thus,

$$E[N_n(j) - M_n(j)] = \lambda_{n-1}E[T_n(j) - W_n(j)]$$
$$= \lambda_{n-1}E[S_n(j)],$$

and substituting this expression into Equation (6), we get

$$E[S_n(j)] \geq \frac{1}{\mu_n} + \frac{\lambda_{n-1}E[S_n(j)]}{\hat{\mu}_{n-1}}. \quad (7)$$

Equation (7) implies that

$$E[S_n(j)] \geq \frac{\frac{1}{\mu_n}}{1 - \frac{\lambda_{n-1}}{\hat{\mu}_{n-1}}} = \frac{1}{(1 - \hat{\rho}_{n-1})\mu_n} = \frac{1}{\hat{\mu}_n}.$$

Consequently, this provides a lower bound for the expected service time that is independent of packet j and thus proves the induction step.

To generalize this result to the case where $k > 0$, we proceed as follows: since during any busy time interval at queue $n-1$ the server at queue n is served at rate k, this is equivalent to supplying the server with full rate capacity for k fraction of time and at zero rate for $(1-k)$ fraction of time during that busy period. Using this equivalence we get,

$$T_n(j) = R_n(j) + V_n(j) + (1-k)\sum_{m=1}^{N_n(j)} S_{n-1}(m);$$

$$W_n(j) = R_n(j) + (1-k)\sum_{m=1}^{M_n(j)} S_{n-1}(m).$$

Again we have $E[V_n(j)] \geq 1/\mu_n$ because a packet can initially be served at the lower rate $k\mu_n$. The result now follows along the same lines as before. ∎

Remark: It is worth noting that Theorem 1 provides a proof of Ross's conjecture [9] for the special case of linear influence queueing networks. This conjecture has already been proven under quite general assumptions for single-server queueing systems (see [6] and references therein), but these results do not appear to be directly applicable to the model considered in this paper.
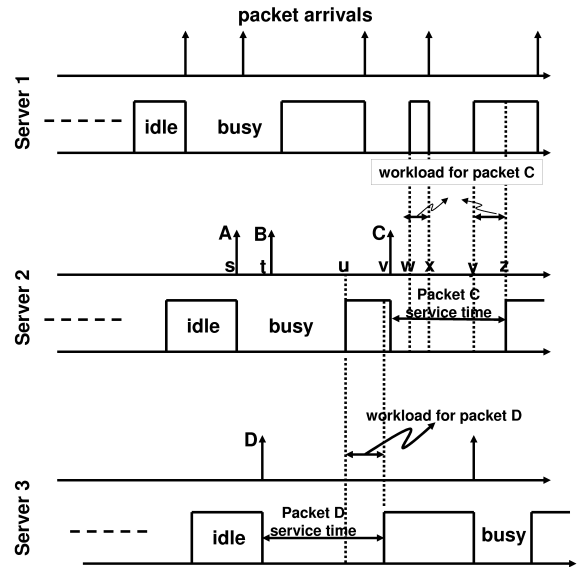

Fig. 3. Depiction of packet arrivals and the corresponding busy and idle periods in a 3-node unidirectional influence queueing network.

Example: To illustrate our results, we compute an exact expression for the queue utilization in a special case of an influence network, that is, a two-node network where $k = 0$. We then compare this expression with the lower bound provided by Theorem 1. This case is similar to a priority queuing system since the influenced queue becomes "blocked" whenever influence exists.

For this specific case, the exact value of ρ_2 can be computed using priority queuing results given in [5]. Let $\pi_{n,m}$ be the joint probability of n users in queue 1 and m users in queue 2. We define the two-dimensional moment-generating function (MGF) of the joint probability as the unilateral z-transform

$$\Pi(z_1, z_2) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \pi_{n,m}z_1^{-n}z_2^{-m}. \quad (8)$$

The MGF of the two-node priority network is given as

$$\Pi(z_1, z_2) = \frac{\mu_2(1 - \rho_1 - \lambda_2/\mu_2)(z_2 - 1)(1 - \eta(z_2)/z_1)^{-1}}{[\mu_1\eta(z_2) - \lambda_1 - \lambda_2(1 - z_2^{-1}) + \mu_2(z_2 - 1)]} \quad (9)$$

where $\eta(z_2)$ is the positive root of the following quadratic equation in the variable x

$$\mu_1 x^2 - [\lambda_1 + \lambda_2(1 - z_2^{-1}) + \mu_1]x + \lambda_1 = 0. \quad (10)$$

The utilization of the second queue is the complementary of the marginal probability that no class 2 users are in the system:

$$\rho_2 = 1 - \lim_{z_2 \to \infty}\lim_{z_1 \to 1}\Pi(z_1, z_2) = \frac{\rho_1 + \lambda_2/\mu_2 + \eta}{1 - \eta}, \quad (11)$$

where

$$\eta = \lim_{z_2 \to \infty}\eta(z_2)$$
$$= \frac{\lambda_1 + \lambda_2 + \mu_1 - \sqrt{(\lambda_1 + \lambda_2 + \mu_1)^2 - 4\lambda_1\mu_1}}{2\mu_1}.$$
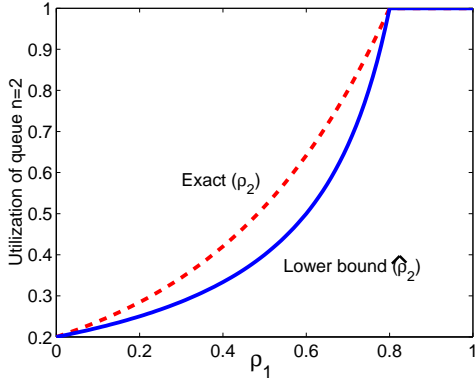
Fig. 4. Utilization of a two-node influence network with $k = 0$.

It is easy to show that the expression provided by Eq. (11) is always larger than $\hat{\rho}_2$, for any values of λ_1, λ_2, μ_1, and μ_2.

Figure 4 compares ρ_2 and $\hat{\rho}_2$ for $\mu_1 = \mu_2 = 1$, $\lambda_2 = 0.2$ and $\lambda_1 \in [0, 1]$. We observe that the lower bound is reasonably close to the actual queue utilization.

IV. PHASE TRANSITION

We now show how the lower bound provided by Theorem 1 leads to a phase transition in the network behavior under certain conditions. We assume the model given in the beginning of Section II-B, with a uniform influence parameter k and packet size parameters $\mu_n = \mu$ for all n.

First, we set the traffic rates λ_n such that all queues $n = 1, 2, 3, ...$ are guaranteed to have a certain minimum utilization $\hat{\rho}_I < 1$ as specified by the lower bound of Theorem 1. Thus,

$$\lambda_1 = \mu\hat{\rho}_I \tag{12}$$
$$\lambda_n = \mu[\hat{\rho}_I - (1-k)(\hat{\rho}_I)^2], \qquad n = 2, 3, ... \tag{13}$$

Now, suppose we increase the traffic only at queue 1, that is, we increase λ_1 but keep the mean arrival rates to all the other queues, λ_n, constant. How will the effect of this increased traffic propagate through the network?

There are two main cases to consider based on the relationship between the parameters $\hat{\rho}_I$ and k. If $\hat{\rho}_I < \frac{k}{1-k}$ it turns out that independent of the value of ρ_1 the downstream queues $n = 2, 3, ...$ will always remain stable (we recall that $\rho_1 = \lambda_1/\mu$). On the other hand, if $\hat{\rho}_I > \frac{k}{1-k}$, then the lower bound on the queue utilization $\hat{\rho}_n$ will experience a phase transition, as ρ_1 is increased. Specifically, we next show that if ρ_1 crosses a certain threshold ρ_1^*, then the value of $\hat{\rho}_n$ suddenly jumps from $\hat{\rho}_I$ to 1, for large values of n. Our simulations in the next section show that a similar phenomenon takes place for the actual queue utilizations ρ_n. Note that the condition $\hat{\rho}_I < 1$ implies that a phase transition can only take place when $k < 0.5$, for the system under consideration.

Theorem 2 *Suppose that a linear influence queuing network is given with mean arrival rates satisfying Eq. (13). Suppose further that $\hat{\rho}_I > \frac{k}{1-k}$ with $k < 0.5$ and define $\rho_1^* = \max(\hat{\rho}_I, \frac{1}{1-k} - \hat{\rho}_I)$. Then, the following result holds:*

$$\lim_{n\to\infty} \hat{\rho}_n = \begin{cases} \hat{\rho}_I & \rho_1 < \rho_1^* \\ 1 & \rho_1 > \rho_1^*. \end{cases} \tag{14}$$

Proof:

We first consider the case $\frac{k}{1-k} < \hat{\rho}_I \leq \frac{1}{2(1-k)}$, for which $\rho_1^* = \frac{1}{1-k} - \hat{\rho}_I$.

Now, assume that $\rho_1 \in [0, \rho_1^* - \alpha]$, where α is a constant such that $0 < \alpha < \frac{1}{1-k} - 2\hat{\rho}_I$. For such value of ρ_1, it is easy to show by induction that $\hat{\rho}_n \in [0, \rho_1^* - \alpha]$. Specifically, assume $\hat{\rho}_{n-1} \in [0, \rho_1^* - \alpha]$. From Eqs. (3) and **??**QN-rates), we have

$$\hat{\rho}_n = \frac{\hat{\rho}_I - (1-k)(\hat{\rho}_I)^2}{1 - (1-k)\hat{\rho}_{n-1}}. \tag{15}$$

We note that $\hat{\rho}_n$ is an increasing function of $\hat{\rho}_{n-1}$. Thus, $\hat{\rho}_n$ achieves its minimum when $\hat{\rho}_{n-1} = 0$ and maximum when $\hat{\rho}_{n-1} = \rho_1^* - \alpha$. Substituting 0 for ρ_{n-1} in Eq. (15), we obtain the following lower bound on $\hat{\rho}_n$:

$$\hat{\rho}_n \geq \hat{\rho}_I - (1-k)(\hat{\rho}_I)^2. \tag{16}$$

It is easy to show that the right-hand side of Eq. (16) is always greater than 0, for any $0 < \rho_I < 1$. Similarly, substituting $\rho_1^* - \alpha$ for ρ_{n-1} in Eq. (15), we obtain the following upper bound on $\hat{\rho}_n$:

$$\hat{\rho}_n \leq \frac{\hat{\rho}_I - (1-k)(\hat{\rho}_I)^2}{1 - (1-k)(\rho_1^* - \alpha)} = \frac{\hat{\rho}_I - (1-k)(\hat{\rho}_I)^2}{(1-k)(\hat{\rho}_I + \alpha)}. \tag{17}$$

Since $\alpha < \frac{1}{1-k} - 2\hat{\rho}_I$, we have, after multiplying both sides of this equation by $\alpha(1-k)$(we remind that $\alpha > 0$) and re-arranging terms,

$$\alpha - 2(1-k)\hat{\rho}_1\alpha - (1-k)\alpha^2 > 0. \tag{18}$$

We now add $\hat{\rho}_I - (1-k)(\hat{\rho}_I)^2$ to both the lhs and rhs of Eq. (18) and obtain, after factoring terms,

$$(\hat{\rho}_I + \alpha) - (1-k)(\hat{\rho}_I + \alpha)^2 > \hat{\rho}_I - (1-k)(\hat{\rho}_I)^2. \tag{19}$$

Dividing both the rhs and lhs of Eq. (19) by $(1-k)(\hat{\rho}_I + \alpha)$, we get

$$\frac{1}{1-k} - (\hat{\rho}_I + \alpha) > \frac{\hat{\rho}_I - (1-k)(\hat{\rho}_I)^2}{(1-k)(\hat{\rho}_I + \alpha)}. \tag{20}$$

We now identify the lhs of Eq. (20) as $\rho_1^* - \alpha$, while the rhs of Eq. (20) is the same as the rhs of Eq. (17), which is an upper bound on $\hat{\rho}_n$. Therefore, $\hat{\rho}_n < \rho_1^* - \alpha$, and the induction is proven.

Our goal now is to show that for any $\rho_1 < \rho_1^*$, the sequence $\hat{\rho}_n$ converges to $\hat{\rho}_I$. First we note, that for any $\rho_1 < \rho_1^*$, there exists some $0 < \alpha < \frac{1}{1-k} - 2\hat{\rho}_I$, such that $\rho_1 \in [0, \rho_1^* - \alpha]$. Consider the "error" expression,

$$e(n) = \hat{\rho}_n - \hat{\rho}_I. \tag{21}$$

We will show that $|e(n)|$ is bounded from above by a geometrically decreasing sequence with ratio factor $\beta < 1$. This will establish that $e(n) \to 0$ and $\hat{\rho}_n \to \hat{\rho}_I$.

It follows by direct substitution of Eq. (15) into Eq. (21) that

$$|e(n)| = \frac{\hat{\rho}_I(1-k)}{1 - (1-k)\hat{\rho}_{n-1}}|e(n-1)| \triangleq C(\hat{\rho}_{n-1})|e(n-1)|.$$

We are left to show that the multiplicative factor satisfies $C(\hat{\rho}_{n-1}) \leq \beta < 1$, where β is a constant. We observe now that $C(\hat{\rho}_{n-1})$ is an increasing function of $\hat{\rho}_{n-1}$. Furthermore,

since $\rho_1 \in [0, \rho_1^* - \alpha]$, we know from the first part of the proof that $\hat{\rho}_{n-1} \le \rho_1^* - \alpha$. Therefore,

$$C(\hat{\rho}_{n-1}) \le C(\rho_1^* - \alpha) = \frac{\hat{\rho}_I}{(\hat{\rho}_I + \alpha)} \triangleq \beta < 1,$$

which completes our proof for the case $\rho_1 < \rho_1^*$.

The case $\rho_1 > \rho_1^*$ is handled in an analogous manner. We first note that if $\rho_1 \in [\rho_1^* + \alpha, 1]$, with $0 < \alpha < 1 - \rho_1^*$, then for any $n \ge 2$, we have $\hat{\rho}_n \in [\rho_1^* + \alpha, 1]$. The proof of convergence of $\hat{\rho}_n$ follows the same line as before, with all the inequalities reversed. Thus, $\hat{\rho}_n$ is lower bounded by the minimum between a geometric sequence with factor ratio $\beta > 1$ and the value of one. Therefore, for some sufficiently large number N_0, $\hat{\rho}_n = 1$, $\forall n \ge N_0$.

The proof for the case $\frac{1}{2(1-k)} \le \hat{\rho}_I < 1$, for which $\rho_1^* = \hat{\rho}_I$, follows the same lines as above. ■

V. SIMULATION RESULTS

In this section, we first present simulations results for linear IQNs, based on the mathematical model introduced in the previous sections of this paper. Next, we present NS simulations demonstrating the existence of the phase transition phenomenon in an IEEE 802.11-based wireless ad-hoc network.

A. Simulation of Linear IQNs

We have conducted Matlab simulations to evaluate the performance of a linear IQN, consisting of $N = 20$ influence queues. For the simulations, packet lengths are generated as exponential random variables with normalized mean $1/\mu = 1$. As specified by the linear IQN model, all queues operate at only two rates: packets in a queue are served at rate of 1 packet per unit of time when the queue's immediate predecessor is idle, and at rate of $k = 0.3$ packet per unit of time when the queue's immediate predecessor is busy. Also as specified earlier, we set all arrival rates as a function to the lower bound of the initial network utilization $\hat{\rho}_I$, and then let $\lambda_1 \to \mu$ (while keeping all other parameters constant) to illustrate the effects of influence propagation. For each value of λ_1, steady-state results were obtained by simulating the network for $T = 2 * 10^5$ units of time and averaging over 5 trials.

Note that from Theorem 2, phase transition may possibly take place only if $\hat{\rho}_I > 0.428$ when $k = 0.3$. In our simulations, we set $\hat{\rho}_I = 0.45$. Figure 5 depicts the simulation results obtained for the queue utilization ρ_n as a function of ρ_1, for queues $n = 2, 5, 10, 20$. These simulation results are compared with the lower bound provided by Theorem 1. In each case, we observe that the qualitative behavior of the simulated results and the lower bound are similar. The lower bound predicts quite well the point at which a sharp increase in the utilization takes place.

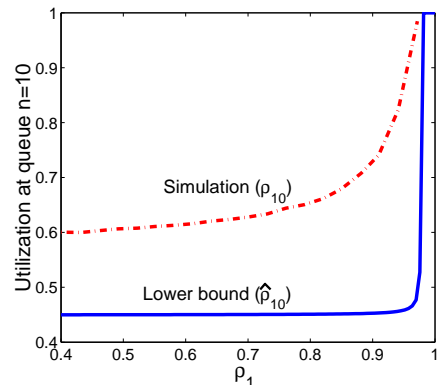### B. Phase Transition in IEEE 802.11 Wireless Networks

In this section, we present simulation results, obtained using the Network Simulator (NS) [1], for a linear IEEE 802.11 wireless ad-hoc network. These results suggest that the influence propagation and phase transition phenomena
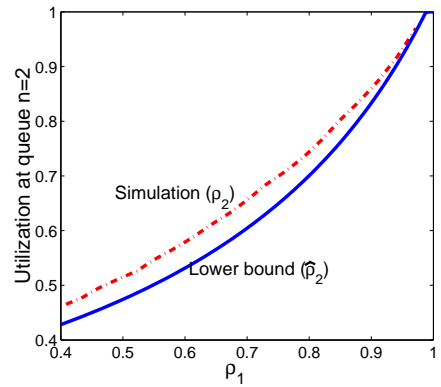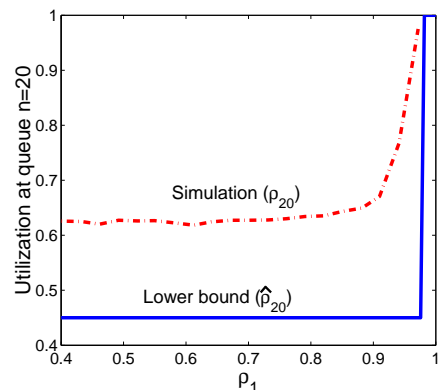

(a) Utilization of queue 2
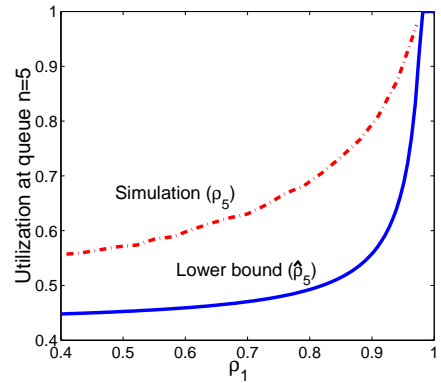
(b) Utilization of queue 5

(c) Utilization of queue 10

(d) Utilization of queue 20

Fig. 5. Utilization of various queues as a function of ρ_1: simulation and lower bound ($k = 0.3$ and $\rho_I = 0.45$).
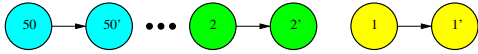
Fig. 6. Simulated IEEE 802.11 network topology

reported in the previous sections may actually take place in real wireless networks. Thus, although our IQN model provides a fairly high-level abstraction of the operation of actual wireless network protocols, it is able to properly capture their qualitative behavior.

Set-up: We consider a linear topology consisting of 50 pairs of nodes, as shown in Fig. 6. Each node n, $n = 1, 2, \ldots, 50$, transmits packets to node n'. In addition, any transmission by node n interferes with packet reception at node $(n+1)'$. For instance, if both nodes 1 and 2 transmit at the same time, then the packet sent by node 2 to node $2'$ is destroyed.

In accordance with the IEEE 802.11 standard, data packet transmissions must be acknowledged [2]. If a node does not a receive an acknowledgement for a packet that it has transmitted, it backoffs for a random period of time and tries again. The maximum number of retransmissions is limited by the "Long Retry Limit" parameter. This parameter is set to 4 in our simulations.

Our statistical assumptions on the network traffic are as follows. At each node n, packets arrive according a Poisson process with rate λ_n. These packets are stored in a very large (lossless) buffer and transmitted in a First-In First-Out (FIFO) order. The size of each packet is fixed and set to 2000 bytes. The data rate of the channel is set to 1 Mb/s. Thus, each node can transmit at most 62.5 packets per second. The total running time of the simulation is 1000 seconds.

The IEEE 802.11 network model under consideration is similar to a linear IQN with influence parameter $k = 0$ (queue n represents the buffer at node n). Note, though, that the linear IQN model does not incorporate back-offs and retransmissions. Thus, we cannot expect both models to produce the same exact results. However, we will show that their qualitative behavior is similar in the sense that the IEEE 802.11 network does experience phase transitions as predicted by the IQN model.

Results: The main performance metric for our simulations is the (channel) utilization at each node of the network. This metric is defined as the fraction of time each node is transmitting (successfully or not) over the course of the entire simulation. The exogenous packet arrival rate λ_n at each node $n \geq 2$ is set to 7 packets per second, while the packet arrival rate λ_1 at the first node is varied from 1 to 62.5 packets per second.

Figure 7 depicts the average utilization at each node of the network, for two different cases: $\lambda_1 = 47$ packets per second and $\lambda_1 = 52$ packets per second (corresponding respectively to utilization $\rho_1 = 0.75$ and $\rho_1 = 0.83$ at node 1). Although the utilization of node 1 varies only by a little amount, the corresponding network behavior is completely different. In the first case the influence of node 1 decays and the network is stable, while in the second case all the nodes
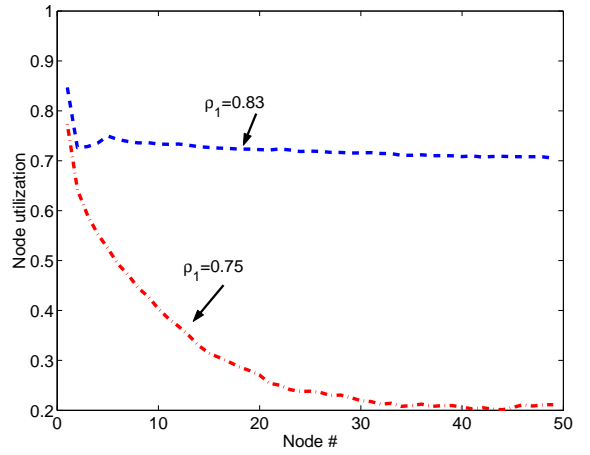

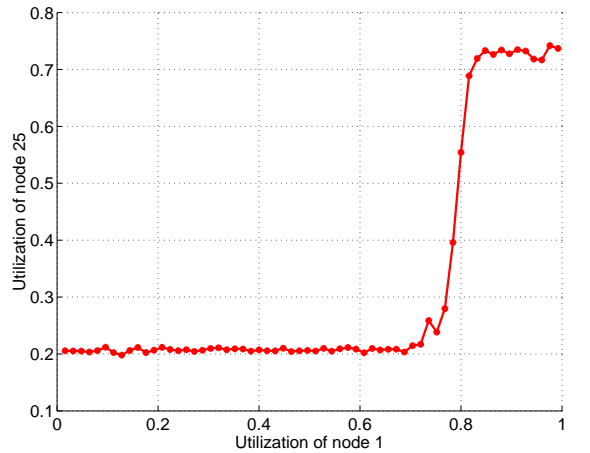Fig. 7. Utilization at each network node for two different values of ρ_1.

Fig. 8. Utilization of node 25 as a function of ρ_1.

in the network become saturated. Note that the maximum utilization in the saturated case is less than 1, due to backoffs and inter-frame spacings required by the IEEE 802.11 standard that prevent nodes from transmitting continuously [2].

The phase transition phenomenon taking place in this IEEE 802.11 network is further illustrated by Fig. 8. This figure depicts the utilization at node $n = 25$ as a function of the load at node 1. We observe that the utilization remains steady around 0.2, as long as $\rho_1 < 0.75$. Then it increases sharply, reaching saturation as ρ_1 exceeds 0.8. This effect is quite dramatic given the fact that node 25 is at a distance of 50 hops from node 1.

For this network, Theorem 2 predicts that phase transition would take place at load no higher than $\rho_1 = 0.87$. From Fig. 8, we observe that it actually takes place at load $\rho_1 \approx 0.8$. This result suggests that Theorem 2 may provide a useful estimate on the critical threshold value of the phase transition.

VI. Concluding Remarks

In this paper, we have studied the impact of distributed carrier sense and rate adaptation mechanisms on the performance and stability of multi-hop wireless networks. We have

shown that while these mechanisms attempt at maximizing the use of local resources, they may have a detrimental impact on global network performance. Specifically, we have introduced a queuing framework, called influence queueing network (IQN), which qualitatively captures interference effects caused by nearby transmitting nodes. Using this framework, we have shown that transmissions by a single node can influence the behavior of the entire network. A surprising result in this context is that the network may exhibit a phase transition, wherein a distant queue may exhibit little change before being forced to instability. The practical implication of these findings has been confirmed by NS simulations of an IEEE 802.11 wireless ad-hoc network.

Furthermore, the analysis and simulations in this work have illustrated how propagating influence can lead to network instability, as a function of the influence parameter k and the overall traffic load on the network. The impact of this trade-off is significant, because it limits how much can be gained by rate control in an ad hoc network, once the minimum available rate drops below a certain threshold with respect to the network load, the network faces a greater chance of instability than if only a single rate were used.

As an example, consider an IEEE 802.11 wireless network which offer data rates of 1, 2, 5.5, and 11 Mb/s. If many such wireless devices were to form a network with a linear influence structure, our results indicate that the low data rates offered would lead to phase transition behavior for sufficiently high levels of traffic. Specifically, if the data rates were limited to either 1 or 11 Mbps (i.e. the binary-rate influence model), phase transition behavior would be predicted for an average network load of only 10%. Furthermore, as shown in the simulated performance, phase transition effects can be quite visible in a linear network with as few as 10 nodes. Consequently, this phenomenon is a realistic concern even for a network with low-to-moderate traffic and a reasonable number of nodes.

This work opens several new directions for research. For instance, a challenging problem is to derive upper bounds on the network utilization that would shed further light into the phase transition phenomenon and complement our existing lower bound. Another interesting research area is in the analysis of non-linear IQNs and bi-directional IQNs where queues can mutually influence each other. Finally, the analysis of IQNs with traffic assumptions other than those used in this paper could also be of interest.

REFERENCES

[1] The network simulator - NS-2. Available on-line at: http://www.isi.edu/nsnam/ns/.

[2] ANSI/IEEE Std 802.11-1999 Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, 1999.

[3] G. Bianchi. Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communication*, 18(3):535–547, 2000.

[4] M. Lacage, M. Manshaei, and T. Turletti. IEEE 802.11 rate adaptation: A practical approach. In *Proceedings of ACM International Symposium on Modeling, Analysis, and Simulation of Wireless and Mobile Systems (MSWiM)*, Venice, October 2004.

[5] R. G. Miller. Priority queues. *Annals of Mathematical Statistics*, 31:86–103, 1960.

[6] M. Miyoshi and T. Rolski. Ross type conjectures on monotonicity of queues. *Australian and New Zealand Journal of Statistics*, 46:121–132, 2004.

[7] M. Neely, E. Modiano, and C. Rohrs. Power allocation and routing in multi-beam satellites with time varying channels. *IEEE Transactions on Networking*, 11(1):138–152, 2003.

[8] S. Ray, J.B. Carruthers, and D. Starobinski. Evaluation of the masked node problem in ad-hoc wireless LANs. *IEEE Transactions on Mobile Computing*, 4(5):430–442, 2005.

[9] S. Ross. Average delays in queues with non-stationary Poisson arrivals. *Journal of Applied Probability*, 15:602–609, 1978.

[10] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Transactions on Automatic Control*, 37(12):1936–1948, Dec 1992.