End-to-End Flow Control for Tandem Packet Radio Networks: Analysis and Simulations

Ibrahim M. Matta
Dept of Computer Science
U of Maryland
College Park, MD 20742, USA
imm@cs.umd.edu

Khalil M. Ahmed
Dept of Computer Science
Faculty of Engineering
U of Alexandria, Egypt
khalil@alex.eg

March 10, 1993

Abstract

We analyze the performance of multihop tandem packet radio networks. In these networks, nodes share a common channel to transmit their packets. Nodes have finite buffers and access the channel according to a collision free channel allocation protocol. We first describe a Markov model of a two–node tandem with end–to–end flow control and obtain analytic expressions for appropriate performance measures. We then study a five–node tandem using simulation. Our results help in understanding the behavior of the channel access protocol, and illustrate the need for the flow control mechanism to avoid congestion and provide fairness in these networks. This appears to be the first analysis of tandem packet radio networks with end–to–end flow control and with finite–buffer nodes accessing the radio channel using a collision free protocol. Suggestions on how to tune the system are discussed, and directions for future research are identified.

Keywords: Radio networks, flow control, performance analysis, modeling, discrete-event simulation.

1 Introduction

Packet radio networks (PR nets) extend packet switching technology into the environment of broadcast radio. They are intended to provide data communications to users located over a broad geographic region, where direct radio or wire connection between the source and destination is not practical [8]. If the two ends of a connection are not within communication range, intermediate radios will rebroadcast (relay) the packet, thus creating a *multihop network*. The design of a packet radio system involves a large number of issues which interact in a very complex way [14]. Many of these concern the radio channel and its use, others concern the operational protocols.

We are concerned with the throughput–delay analysis in multihop PR nets. This is not an easy task, even for simple systems, due to the large amount of state information needed in the state description. A number of studies of throughput-delay in these networks (e.g. [24] [25]) have appeared for some channel access protocols, mainly slotted ALOHA [1]. Tandem packet radio systems have received some attention in the literature [11]. The interesting thing about a tandem configuration is that it is a model of a simple but typical multihop system. Sidi [21] studied two tandem networks (see Figure 1). The first is a tandem where packets arrive to the network only at the top node N. The second is a tandem where packets arrive at all the nodes. All the traffic was assumed to be destined for the station. It was also assumed that the nodes of the tandem have infinite buffers, and they share a common radio channel for transmission of data packets. The channel access scheme was taken to be slotted ALOHA. For these networks, Sidi obtained analytic expressions for the average queue length, the average time delay at each node, and the total average delay in the network.

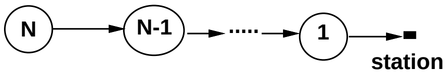

Figure 1: Tandem packet radio network with N nodes.

In this paper, we further investigate tandem packet radio networks. We assume that nodes have *finite buffers* and access the common radio channel using a *slotted collision free* allocation protocol. To our knowledge, this is the first analysis of PR nets with a collision free channel access scheme, leading to multidimensional Markov chain model with product form solution. We evaluate the performance of the network with and without end–to–end flow control. We show that the adoption of the flow control mechanism improves performance, and we give suggestions on how to tune the system.

2 Overview

A schematic figure of the network we consider is depicted in Figure 2. The network consists of N nodes having finite buffers. The nodes transmit data over a common shared radio channel. Fixed-length data packets enter the system at the nodes from corresponding sources. Time is divided into slots of size corresponding to the transmission time of a packet and transmissions are started only at the beginning of a slot.

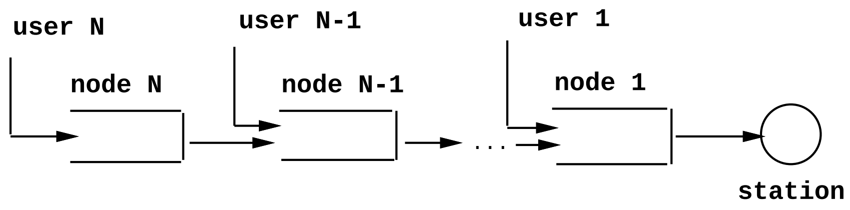

Figure 2: Our target system.

The nodes of the tandem have radio transmitters with omnidirectional antennas and their transmission range is such that a transmission at node i ($2 \leq i \leq N-1$) can be heard only by nodes $(i-1)$ and $(i+1)$. Nodes 1 and N can be heard at nodes 2 and $N-1$ respectively. The final destination of all packets is a station (Fig. 2) that receives the packets transmitted by node 1, and packets entering the network at any node i are forwarded via nodes $i-1, i-2, \ldots, 1$ until they reach the station. We assume that a node can transmit only one packet in a given slot according to a FCFS policy.

We adopt a collision free channel allocation protocol. This approach proposed in [3] for designing data link protocols in the radio environment is based on the allocation of transmission rights to nodes at the beginning of each time slot. This allocation provides each node having the transmission right with a guaranteed, collision free access to the broadcast channel, i.e. its transmission will be correctly received by all its neighbors. The protocol adopted is completely distributed involving only the locality of each node.

We assume that a radio node can receive only a single transmission at a time and cannot transmit and receive simultaneously. To ensure that a node v not only receives but also correctly interprets a transmission initiated by its neighbor, the following conditions must be satisfied for every node v :

Condition C1: A neighbor of v do not transmit in the same slot as v.

Condition C2: Two neighbors of v do not transmit in the same slot.

The two conditions provide collision free commmunication. Note that condition C2 implies that every node u and any of its 2-hop neighbors do not transmit in the same slot.

The outline of the distributed protocol that ensures the above conditions C1 and C2 can be given as follows. We assume that nodes have distinct identities and are aware of the identities of their neighbors. Nodes entering the algorithm exchange list of neighbors information with their neighbors. At this point each node becomes aware of its network locality consisting of all nodes one and two hops away (2-hop neighbors). Node 1 whose identity is minimal (i.e. closer to the station) in its two-hop locality can, at this point, assign itself the transmission right if the proper conditions for transmitting successfully (i.e. its queue is not empty and the station is idle) are satisfied. The assignment information is broadcasted to its neighbors who forward it to nodes which are two hops away from the assigning node. At this point the algorithm proceeds with the next node which is node 2 if node 1 has not assigned itself a transmission right or node 4 if node 1 has assigned itself a transmission right. The algorithm terminates when every node has or has not assigned itself a transmission right.

This approach is algorithmically more complex than slotted ALOHA adopted in [21] since it invests considerably larger efforts into providing guaranteed collision free access to the radio channel which facilitates broadcasting [3].

The two conditions C1 and C2, and the feature of finite buffers at each node imply several facts: (i) The transmissions of node 1 are always successful if the station is idle since no other node can interfere; (ii) The transmissions of node 2 are successful only if node 1 is not transmitting at the same time, and there is a space in queue 1; (iii) The transmissions of node i ($3 \leq i \leq N$) are successful only if both nodes $(i-1)$ and $(i-2)$ are not transmitting, and there is a space in queue $(i-1)$. From this discussion, it is clear that the status of a node i affects the behavior of other nearby nodes.

The rest of the paper is organized as follows. Section 3 discusses a queueing model for a two-node tandem packet radio network which incorporates the collision free access protocol described above and a flow control mechanism. Section 4 extends the two-node tandem network to more than three nodes in order to study the *spatial reuse* phenomena of the radio channel due to the possible simultaneous use of the common channel by nodes $i, i+3, i+6, \ldots$, and exploit the special features of these networks. Section 5 concludes the paper.

3 Analysis of the Two–Node Tandem

In this section, we analyze a Markov model of a two–node tandem packet radio network. We assume that packets arrive to both nodes from corresponding sources. We obtain analytic expressions for the throughput at each node, the average time delay at each node, the total average delay in the network, and the total throughput of the network. We attempt to reduce congestion occuring at high traffic by imposing a limit on the number of messages each user may have anywhere along the tandem link. This is called an *end-to-end control* [18, 5] scheme. From now on, $EE(j)$ will denote that the control strategy is applied and that for each user, no more than j messages may exist in the system. No new messages are allowed to enter the network until some of the user's previous messages have cleared out of the network. Next, we present numerical results that illustrate the tradeoffs among offered load, buffer capacity, and end-to-end control parameter with regard to their impact on network throughput and delay. Finally, simulation results are compared to those obtained from analysis. This comparison manifests good agreement between both results.

3.1 Markov Model

The queueing network representation shown in Figure 3 is adopted with the addition of an end-to-end scheme for congestion control.

Referring to Figure 3, it is assumed that λ_1 and λ_2 represent the average load, in packets/slot,

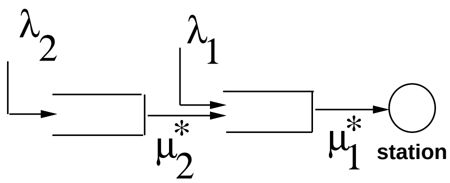

Figure 3: Two-node tandem.

arriving at nodes 1 and 2 respectively. μ_2^* represents the rate of scheduled[1] transmissions from node 2 to node 1, in packets/slot. μ_1^* represents the rate of scheduled transmissions from node 1 to the station, in packets/slot. We assume Poisson arrivals and scheduled transmissions. The processing time at the station is negligible. Each node has a pool of L buffers. EE[2] is the end-to-end control parameter at each node.

For the moment, let the effective transmission rates of the queues μ_1^* and μ_2^* be known quantities. (The determination of these transmission rates is discussed later.)

Let n_{22} denote the number of packets, generated by user 2, at node 2. Let n_{11} and n_{21} denote the number of packets at node 1, generated by user 1 and user 2, respectively. We characterize the state of the system by the state variable (n_{22}, n_{21}, n_{11}). Because of limited buffer sizes and with end–to–end flow control, the state space A is

$$\begin{aligned} A &= \{(n_{22}, n_{21}, n_{11}) : 0 \leq n_{22} \leq L \textit{ and } 0 \leq n_{22} + n_{21} \leq EE \\ &\quad \textit{and } 0 \leq n_{21} + n_{11} \leq L \textit{ and } 0 \leq n_{11} \leq EE\} \end{aligned}$$

For a FCFS discipline, at node 1, the probability that a first user's packet is in service (being transmitted to the station) is $n_{11}/(n_{21} + n_{11})$, while the probability that a second user's packet is in service is $n_{21}/(n_{21} + n_{11})$. The rate of service of first user's packets is thus $\mu_1^*\, n_{11}/(n_{21} + n_{11})$, while the rate of service of second user's packets is $\mu_1^*\, n_{21}/(n_{21} + n_{11})$. Thus our model is a queueing system with state–dependent service.

The transition–rate diagram in the interior of the state space A is shown in Figure 4. We obtain the following balance equation:

[1] Scheduling is done by the channel access protocol.
[2] Here we simply use the notation EE instead of $EE(j)$ for some fixed j.

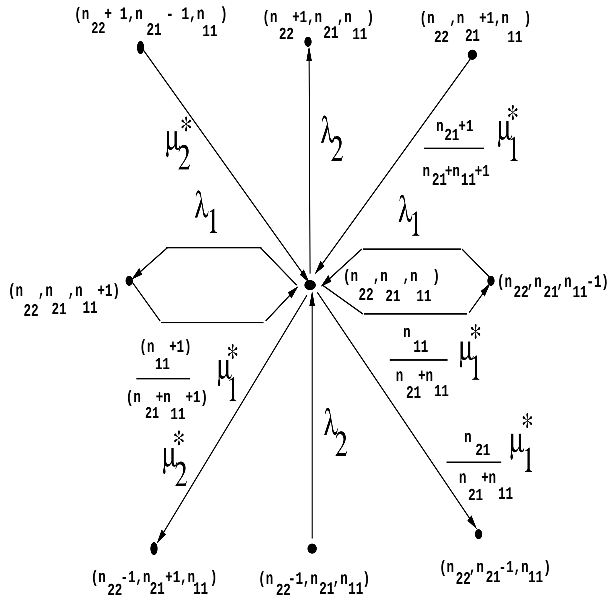

Figure 4: Transition rates, interior region, Fig. 3.

$$
\begin{aligned}
&(\lambda_1 + \lambda_2 + \mu_1^* + \mu_2^*) P(n_{22}, n_{21}, n_{11}) = \\
&\qquad \lambda_2 P(n_{22}-1, n_{21}, n_{11}) + \lambda_1 P(n_{22}, n_{21}, n_{11}-1) + \mu_2^* P(n_{22}+1, n_{21}-1, n_{11}) + \\
&\qquad \frac{(n_{21}+1)\mu_1^*}{(n_{21}+n_{11}+1)} P(n_{22}, n_{21}+1, n_{11}) + \frac{(n_{11}+1)\mu_1^*}{(n_{21}+n_{11}+1)} P(n_{22}, n_{21}, n_{11}+1)
\end{aligned} \tag{1}
$$

We leave it to the reader to derive the equations in the boundary cases such as $n_{22} = n_{21} = n_{11} = 0$. It is easy to verify by substitution that Equation 1 is satisfied by the following product-form solution:

$$
P(n_{22}, n_{21}, n_{11}) \;=\; \frac{1}{S} \left(\frac{\lambda_2}{\mu_2^*}\right)^{n_{22}} \left(\frac{\lambda_2}{\mu_1^*}\right)^{n_{21}} \left(\frac{\lambda_1}{\mu_1^*}\right)^{n_{11}} \frac{(n_{21}+n_{11})!}{n_{21}!\; n_{11}!} \tag{2}
$$

where S is the normalization constant.

$$
S \;=\; \sum_{(n_{22}, n_{21}, n_{11}) \epsilon A} \left(\frac{\lambda_2}{\mu_2^*}\right)^{n_{22}} \left(\frac{\lambda_2}{\mu_1^*}\right)^{n_{21}} \left(\frac{\lambda_1}{\mu_1^*}\right)^{n_{11}} \frac{(n_{21}+n_{11})!}{n_{21}!\; n_{11}!} \tag{3}
$$

The only remaining question is that of calculating the effective transmission rates μ_1^* and μ_2^*. These rates depend on the channel access protocol. Both nodes can actually transmit one packet/slot. For the slotted collision free protocol adopted, node 2 transmits only if node 1 is silent and there is a space in queue 1. Since we assume zero processing time at the station, then node 2 is blocked as long as node 1 is not empty. Thus,

$$\begin{aligned} \mu_1^* &= 1 \; packet/slot \\ \mu_2^* &= 1 \times P\;[n_{21} + n_{11} \;=\; 0] \; packet/slot. \end{aligned} \tag{4}$$

We note that μ_2^* depends on the probability that node 1 is empty, which in turn depends on the state probability $P(n_{22}, n_{21}, n_{11})$, which is itself a function of μ_2^* (see equation 2). We obtain μ_2^* iteratively as follows. We choose an initial value $\mu_2^*(0)$, compute the normalization constant S and the state probabilities $P(n_{22}, n_{21}, n_{11})$'s. We define $\mu_2^*(1) \;=\; P\;[n_{21} + n_{11} \;=\; 0]$ and repeat the process k times until $|\mu_2^*(k) \;-\; \mu_2^*(k-1)| < \epsilon$ for some $\epsilon > 0$.

We now define the performance measures of interest. These include the input-packet blocking probability at each of the nodes for end-to-end control and finite buffers. At node 2, the input packets are blocked whenever $n_{22} = L$ or $n_{22} + n_{21} = EE$. At node 1, the input packets are blocked whenever $n_{21} + n_{11} = L$ or $n_{11} = EE$. Therefore,

$$P_{B_1} \;=\; 1 - \sum_{\substack{0 \le n_{11} < EE \\ 0 \le n_{21} + n_{11} < L}} P(n_{22}, n_{21}, n_{11}) \tag{5}$$

$$P_{B_2} \;=\; 1 - \sum_{\substack{0 \le n_{22} + n_{21} < EE \\ 0 \le n_{22} < L}} P(n_{22}, n_{21}, n_{11}) \tag{6}$$

The throughput for the first user is $\lambda_1(1 - P_{B_1})$. The throughput for the second user is $\lambda_2(1 - P_{B_2})$. The total network throughput is $\lambda_1(1 - P_{B_1}) + \lambda_2(1 - P_{B_2})$.

Using Little's formula [15], the average time delay T_2 for the second user is

$$T_2 = \frac{E[n_{22}]}{\lambda_2(1 - P_{B_2})} + \frac{E[n_{21} + n_{11}]}{\lambda_1(1 - P_{B_1}) + \lambda_2(1 - P_{B_2})} \tag{7}$$

where $E[n_{22}]$ and $E[n_{21}+n_{11}]$ are the expected number of packets queued at nodes 2 and 1, respectively. Both $E[n_{22}]$ and $E[n_{21}+n_{11}]$ may be obtained from the state probabilities $P(n_{22}, n_{21}, n_{11})$.

The average time delay T_1 for the first user is

$$T_1 = \frac{E[n_{21}+n_{11}]}{\lambda_1(1-P_{B_1}) + \lambda_2(1-P_{B_2})} \tag{8}$$

The total average delay T in the network is

$$T = \frac{E[n_{22}] + E[n_{21}+n_{11}]}{\lambda_1(1-P_{B_1}) + \lambda_2(1-P_{B_2})} \tag{9}$$

3.2 Numerical Results

We now present numerical results. We first consider the uncontrolled network then the network with end–to–end flow control. Throughout the rest of the paper, the arrival rates at all nodes are identical. We denote $\lambda = \lambda_1 = \lambda_2$.

Uncontrolled network

Figure 5 shows the throughput as a function of the offered load λ for $L = 3$. We observe that as λ increases, the network throughput increases linearly with λ, then turns down and decreases. Notice that the second user throughput decreases rapidly to zero at high load. Figure 6 shows the throughput as a function of the buffer size L of each node for $\lambda = 0.6$ packet/slot. We observe a degradation in the second user throughput as L is increased while the first user throughput increases. Increasing either L or λ increases the net arrival rate for the first user while at the same time decreasing the probability that node 2 can transmit. Consequently, the input-packet blocking probability at node 2 increases and hence the second user throughput decreases.

Controlled network

Figures 7, 8 show the throughput as a function of the end-to-end control parameter for $L = 3$ and $\lambda = 0.8$ and 1.2 packet(s)/slot respectively. The control parameter is shown varying from 1 to 5. Note that the network throughput increases to a maximum as the control parameter is increased then turns down and decreases. Note also that for a larger offered load, the maximum achievable network throughput is larger, and the maximum second user throughput occurs at a smaller value of EE. Also, observe that both the maximum first user throughput and the maximum network throughput occur at $EE = L$ since increasing EE from 1 to L reduces the input-packet blocking probability at node 1. Similar results were obtained [16] for larger L and different λ. The maximum achievable network throughput is found to be slightly larger for a larger L for each λ. This indicates that the network performance is not dictated by the storage capacity at the nodes (i.e. not storage bound). However, the channel access protocol is the dominant factor in determining the performance of such networks.

The second user experiences more time delay than the first user. The second user delay for $L = 3$ and $\lambda = 1.2$ packets/slot appears in Figure 9. From both Figures 8 and 9, we observe that the time-delay performance is much more sensitive to the choice of EE than was the throughput performance. Thus reducing EE from 2 to 1 reduces the time delay by a factor of 2.4, while the throughput of Figure 8 is affected much less; the network throughput is increased only by a factor of 1.2 by increasing EE from 1 to 2. Clearly, there is a tradeoff between throughput and delay. Typically, this tradeoff is captured in a measure called "power" [10]. We will look at this in section 4.

Figure 10 shows the throughput as a function of the offered load λ for $L = 3$ and $EE(1)$. Figure 11 shows the total average time delay in the network as a function of the offered load for $L = 3$ in the case of $EE(1)$ and when no control strategy applied. Notice that with end-to-end control, the network throughput does not turn down as the offered load is increased to infinity. In addition, degradation in the second user throughput is considerably reduced, and the total average time delay is significantly lower. These results indicate that end-to-end controls are a simple and effective means to control congestion in these networks. Section 4 discusses how this flow control scheme can be properly designed to give the best performance.

3.3 Simulation Results

The network shown in Figure 3 is simulated to verify our analytical analysis. The slotted collision free channel allocation protocol and the end-to-end control scheme are incorporated in the simulation. All messages generated are single packets of fixed size. The transmission time of a packet is equal to the slot size. New messages are generated according to a Poisson process.

Figure 12 shows the throughput versus EE control parameter. We note the following discrepancy. The network throughput and the first user throughput given by analysis with negative slope are not realizable in simulation. Simulation shows that the throughput saturates at $EE(L)$. This difference comes from the Poisson assumption used in formulating the analytical model since the transmission times in the simulator are constant (equal to the size of the slot). (Queueing systems characterized by a larger variance typically have worse performance [13].)

On the other hand, our simulation results [16] support conclusions drawn from our analytic results. In particular, the network is not storage bound, and the use of the end-to-end flow control scheme is necessary to avoid second user throughput degradation, and a large total average time delay.

4 Simulation of the Five-Node Tandem

In this section, we analyze the performance of a five-node network using simulation. We illustrate the spatial reuse phenomena of the radio channel due to the possible simultaneous use of the radio channel by nodes one and four or by nodes two and five. We show how congestion is reduced by invoking the end-to-end control scheme, and how this scheme prevents unfairness in this multiuser environment in spite of its negative effect on overall throughput efficiency.

4.1 Simulation Model and Assumptions

The simulation model developed incorporates the slotted collision free channel access protocol and the end-to-end control strategy as explained previously. All messages generated in the simulation are single packets of fixed size equal to the slot size. New messages are generated according to a Poisson process. The load on all nodes is assumed to be homogeneous. The processing time at the station follows an exponential distribution with mean 0.1 slot. The performance measures of interest are throughput and delay.

4.2 Simulation Results

The classical method of *independent replications* is used to estimate the simulation performance measures [22]. The length of the simulation runs was chosen long enough to ensure steady-state behavior. This behavior is detected when the average arrival rate into each node is the same as the departure rate from that node.

Simulation results were compared with those obtained by Sidi [21] for the case where packets arrive only to the top node (i.e. node 5), assuming zero processing time at the station, infinite queues at the nodes, and an uncontrolled network. In this case, both the slotted ALOHA protocol (where each node transmits in a slot whenever it has a packet ready for transmission) adopted by Sidi and the slotted collision free protocol adopted by us behave similarly. Our simulation results agreed exactly with the analytical results found in [21].

Uncontrolled network

First, the performance of the uncontrolled network is investigated. We assume all nodes to have the same maximum queue size $Q = 10$. Figure 13 shows the throughput versus offered load. Users 2, 3, 4 and 5 suffer from congestion. It is clear that the action taken by each node depends on the actions taken by neighboring nodes. At light traffic, nodes 1 and 2 transmit alternately. Node 1 transmits when the station is idle, and node 2 transmits when node 1 is waiting for the station to be idle. The fourth node has the chance to transmit with the first node at the same time, while the fifth node has the chance to transmit with the second node at the same time. This degrades significantly the throughput of the third user. Note that the throughput of the fifth user is better than the fourth since no transit messages enter its queue. At heavy traffic, when the queues become full, the throughput degrades for all users except the first one.

Figure 14 shows the delays experienced by each user. It is evident that the first user is the one who experiences the minimum delay since he is the closest one to the station. We observe a considerable increase in time delay at the third node. This increase is produced by the first and second nodes which transmit alternately based on the station's status (busy or idle). When the average interarrival time at each node increases beyond 15 slots, the third, fourth and fifth users suffer from a large increase in time delay. It is this increased time delay that manifests itself as *congestion*. From Figure 13, we notice that at this offered load, the third user throughput turns down and decreases towards zero. The network throughput at this load is our network capacity, and is equal to 0.33 packet/slot.

Intuitively, a node should carry as much traffic from its source as from other nodes (i.e. acting as a repeater). We define fairness as "the throughput at each node decreases as the distance from the station increases. In addition, both the throughput and the average time delays for all users are kept to tolerable proportions [18]." Observe that fairness is violated in our uncontrolled network.

Figures 15 and 16 show that the fifth user time delay in the network is greater for a larger $Q = 200$. In addition, the network throughput is slightly greater. The delay increases with the increase in the offered load, but the network capacity, defined as the network throughput at which the delay becomes theoretically infinite, is almost the same for both considered queue sizes. As we concluded from our analytical results, the network is not storage bound, and increasing the storage capacity at the nodes does not improve the performance of such networks. In fact, the network performance is dictated essentially by the behavior of the channel access protocol.

Controlled network

We now evaluate the performance of the network with end–to–end flow control. We assume that $Q = 10$. Figure 17 shows the total average time delay versus network throughput for different values of offered load. We observe that with tighter control, i.e. smaller values of EE, the delay decreases significantly. Besides the throughput decreases due to the increase in the input-packet blocking probability. Clearly, there is a tradeoff between throughput and delay. A more compact measure of combined throughput and delay performance is offered by the concept of "power" [10]. Power is defined as the ratio of throughput over delay multiplied by one minus the blocking probability. One may choose EE that maximizes the value of power. Figure 18 shows the power versus the control parameter EE for different values of offered load. Note that at heavy traffic, the power is maximum at $EE(1)$.

Figure 19 shows that the delays with $EE(1)$ adopted are much lower. Figure 20 shows the throughput versus offered load for the controlled $EE(1)$ case. Congestion is reduced and even eliminated. In addition, fairness is provided. The throughput decreases as the distance (number of hops) from the station increases. Now, the throughput of each user is limited or, equivalently, he is blocked from entering a message until his previous message has cleared out of the system. The congestion produced by the first and second users, and experienced by the third user is reduced. Thus the probability that the third node can transmit is increased. This in turn decreases the probability that nodes 4 and 5 transmit. Moreover, the congestion experienced by the fourth user as a result of the fifth user increase in traffic rate is reduced. So the control mechanism really reduces congestion throughout the network and keeps the throughput of all users to tolerable proportions. Figure 20 reveals a decrease in the first user's throughput with the control scheme adopted, but shows only a decrease in the other users' throughput until a certain traffic intensity after which the throughput increases over the throughput exhibited in the uncontrolled case. We observe that the user experiencing more congestion is the one who benefits most from the control mechanism. This leads us to think of an adaptive $EE(1)$ scheme to be invoked only when the traffic intensity exceeds the network capacity. Figure 21 shows that the network throughput decreases with the control scheme adopted, but congestion is significantly limited in the network and fairness is provided.

5 Conclusions

Two multihop tandem packet radio networks have been studied. Nodes access the channel using a slotted collision free protocol. The end-to-end control strategy is invoked to limit congestion.

For the two-node system, we have developed a solution methodology that is of interest. We have considered a five-node network and have shown that there is a strict difference between the five-node system and a system with up to three nodes, due to the possible simultaneous use of the radio channel by nodes one and four or by nodes two and five. The flow control mechanism adopted prevents throughput degradation and loss of efficiency due to overload and provides fairness in this multi-user environment.

We can identify several problems for future study. A typical design problem might be: Given a tandem network, what is the optimum EE parameter, at each node, that can be invoked to maximize network throughput and minimize total average time delay (within given fairness conditions). These optimum values of EE may be fixed, or may be dynamically adjusted based on traffic conditions. From Figure 17, when the average interarrival time at each node is 15 slots, one may set EE at each node to 2 based on the following reasons: (i) the network throughput is greater than its value with $EE(1)$; (ii) the total average time delay is only increased by a factor of 1.3 when increasing EE from 1 to 2; and (iii) unfairness results if $EE(3)$ is selected as shown in Figure 22. When the average interarrival time at each node is 5 slots, one may set EE to 1 since increasing EE from 1 to 2 increases the network throughput by a factor of 1.1, whereas the total average time delay is affected much more (increased by a factor of 2.2). So $EE(2)$ may be rejected to minimize delay. From this discussion, one may select $EE(2)$ at light load, and $EE(1)$ at high load. The tradeoff between delay and throughput can be captured in the concept of power; this leads to system designs whose optimum operating point is easily found and which corresponds to the operating point one would intuitively choose as shown in Figure 18. Notice that in all cases, the optimum value of EE is small. Also, the adoption of the end-to-end control scheme suggests that, for an N node system, buffers of sizes $N \times EE, (N-1) \times EE, \ldots, 2 \times EE, EE$ are to be designed at nodes $1, 2, \ldots, N-1, N$ respectively. An alternative way to capture the delay-throughput tradeoff is to impose thresholds on the minimum throughput and maximum delays that can be tolerated, and choose the value of EE that is optimum and which satifies these constraints. The analytical attachement of the dynamic control of EE and the study of much more general definitions of fairness and power suggest challenging research problems.

Acknowledgment

We would like to thank Jean-Chrysostome Bolot for his helpful comments and valuable suggestions.

References

[1] N. Abramson. The throughput of packet broadcast channels. *IEEE Transactions on Communications*, COM-25, no.1:117–128, January 1977.

[2] V. G. Cerf and R. E. Kahn. A protocol for packet network interconnection. *IEEE Transactions on Communications*, COM-22, no.5:637–648, May 1974.

[3] I. Chlamtac and S. Pinter. Distributed nodes organization algorithm for channel access in a multihop dynamic radio network. *IEEE Transactions on Computers*, C-36, no.6:728–737, June 1987.

[4] D. W. Davies. The control of congestion in packet-switching networks. *IEEE Transactions on Communications*, COM-20, no.3:546–550, June 1972.

[5] M. Gerla and L. Kleinrock. Flow control: a comparative survey. *IEEE Transactions on Communications*, COM-28, no.4:553–574, April 1980.

[6] I. Gitman, R. M. Van Slyke, and H. Frank. Routing in packet-switching broadcast radio networks. *IEEE Transactions on Communications*, COM-24:926–930, August 1976.

[7] R. Kahn, S. Gronemeyer, J. Burchfiel, and R. Kunzelman. Advances in packet radio technology. In *Proceedings of the IEEE*, volume 66, no.11, pages 1468–1496, November 1978.

[8] R. E. Kahn. The organization of computer resourses into a packet radio network. *IEEE Transactions on Communications*, COM-25:169–178, January 1977.

[9] R. E. Kahn and W. R. Crowther. Flow control in a resource-sharing computer network. *IEEE Transactions on Communications*, COM-20, no.3:539–547, June 1972.

[10] L. Kleinrock. Power and deterministic rules of thumb for probabilistic problems in computer communications. In *Proc. Int. Conf. Commun.*, June 1979.

[11] L. Kleinrock and J. Silvester. On the behavior of multi-hop packet radio networks. In *Proceedings of the IEEE*, pages 156–167, January 1987.

[12] L. Kleinrock and F. A. Tobagi. Packet switching in radio channels: Part i-carrier sense multiple access modes and their throughput delay characteristics. *IEEE Transactions on Communications*, COM-23:1400–1416, December 1975.

[13] S. Lam and M. Reiser. Congestion control of store-and-forward networks by input buffer limits - an analysis. *IEEE Transactions on Communications*, COM-27, no.1:127–133, January 1979.

[14] B. M. Leiner, D. L. Nielson, and F. A. Tobagi. Issues in packet radio network design. In *Proceedings of the IEEE*, pages 6–20, January 1987.

[15] J. D. C. Little. *A proof for the queueing formula* $L = \lambda W$*, Op. Res., vol.9*, pages 383–387. 1961.

[16] I. M. Matta. Performance evaluation of tandem packet radio networks. Master's thesis, Computer Science Department, Alexandria University, Alexandria, Egypt, July 1989.

[17] M. F. Morris and P. F. Roth. *Computer Performance Evaluation*, chapter 7. Van Nostrand Reinhold Company, 1982.

[18] M. C. Pennotti and M. Schwartz. Congestion control in store-and-forward tandem links. *IEEE Transactions on Communications*, COM-23, no.12:1434–1443, December 1975.

[19] M. Schwartz. *Computer Communication Network Design and Analysis*, chapter 13. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1977.

[20] M. Schwartz. *Telecommunication Networks: Protocols, Modeling and Analysis*, chapter 5. Addison–Wesley publishing Co., 1987.

[21] M. Sidi. Tandem packet radio queueing systems. *IEEE Transactions on Communications*, COM-35, no.2:246–248, February 1987.

[22] H. A. Taha. *Operations research : an introduction*. Macmillan publishing Co., second edition, 1976.

[23] F. Tobagi. Modeling and performance analysis of multihop packet radio networks. In *Proceedings of the IEEE*, pages 135–155, January 1987.

[24] F. A. Tobagi. Analysis of a two-hop centralized packet radio network - part I: Slotted ALOHA. *IEEE Transactions on Communications*, COM-28, no.2:196–207, February 1980.

[25] F. A. Tobagi. Analysis of a two-hop centralized packet radio network - part II : Carrier sense multiple access. *IEEE Transactions on Communications*, COM-28, no.2:208–216, February 1980.

[26] F. A. Tobagi. Multi-access protocols in packet communication systems. *IEEE Transactions on Communications*, COM-28:468–488, April 1980.

[27] F. A. Tobagi and L. Kleinrock. Packet switching in radio broadcast channels: Part II-the hidden terminal problem in carrier sense multiple-access and the busy tone solution. *IEEE Transactions on Communications*, COM-23, no.12:1417–1433, December 1975.

[28] H. Zimmerman. OSI reference model-the ISO model of architecture for open systems interconnection. *IEEE Transactions on Communications*, COM-28:425–432, April 1980.

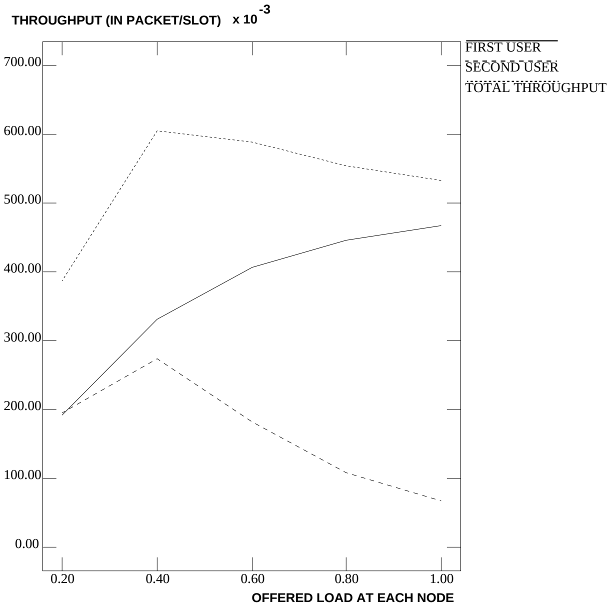

Figure 5: Throughput versus offered load ($L = 3$, uncontrolled).

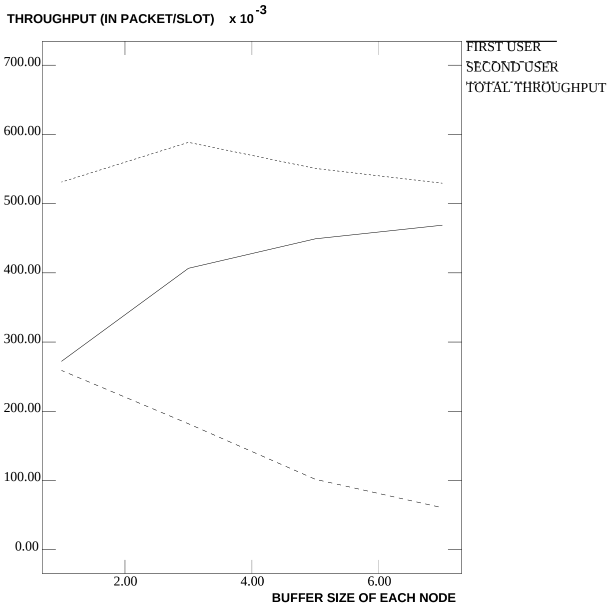

Figure 6: Throughput versus buffer size ($\lambda = 0.6$ packet/slot, uncontrolled).

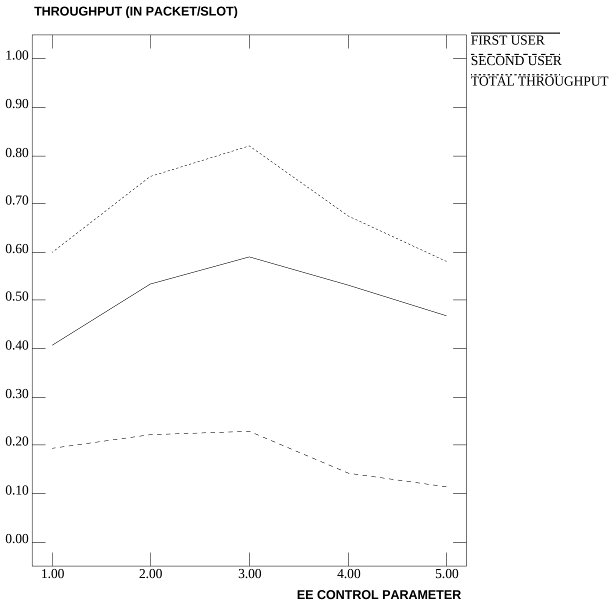

Figure 7: Throughput versus EE ($L = 3$, $\lambda = 0.8$ packet/slot).

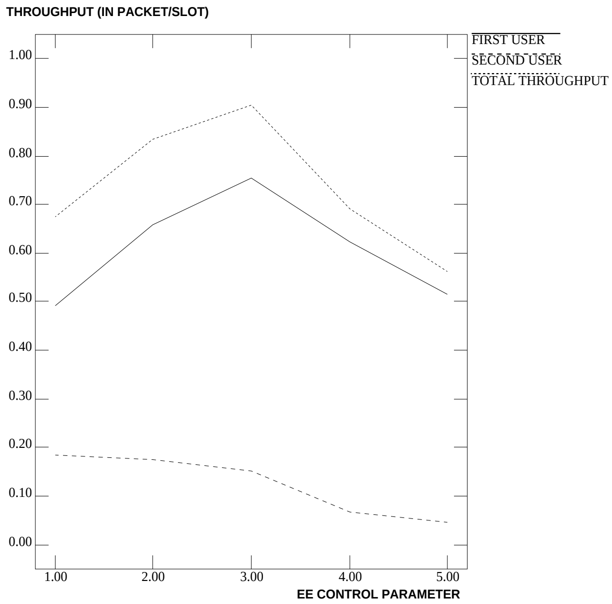

Figure 8: Throughput versus EE ($L = 3$, $\lambda = 1.2$ packet/slot).

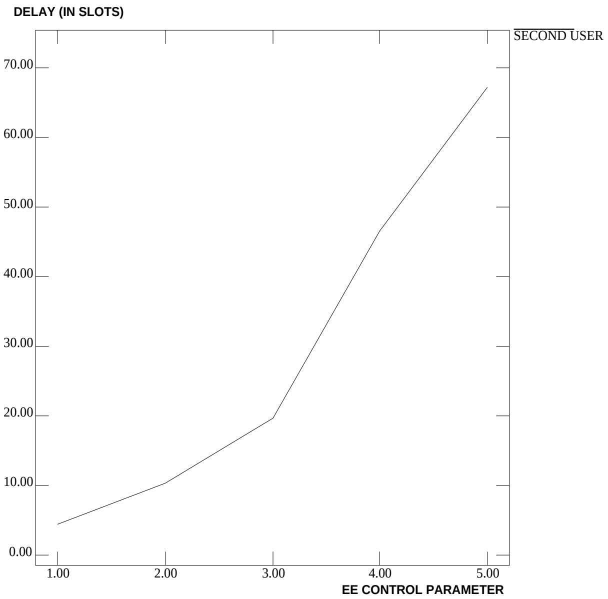

Figure 9: Second user delay versus EE ($L = 3$, $\lambda = \lambda_1 = \lambda_2 = 1.2$ packet/slot).

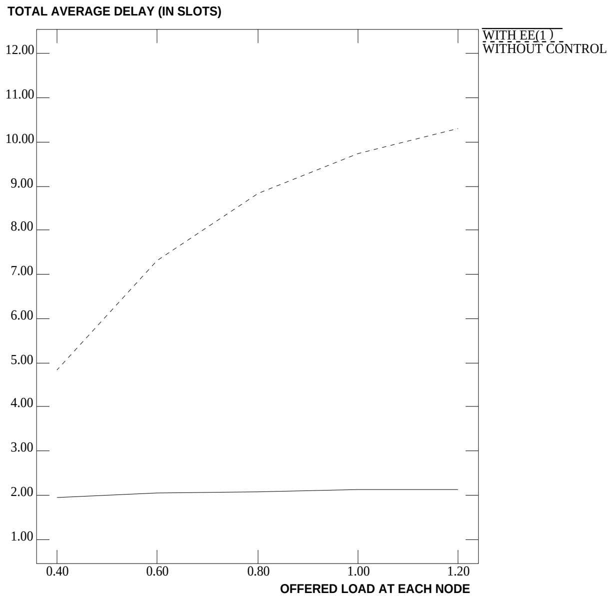

Figure 11: Total average time delay versus λ ($L = 3$).

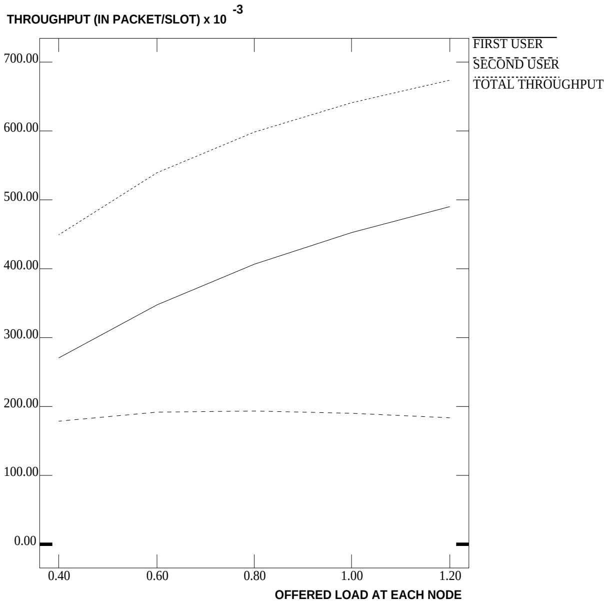

Figure 10: Throughput versus λ ($L = 3$, with $EE(1)$).

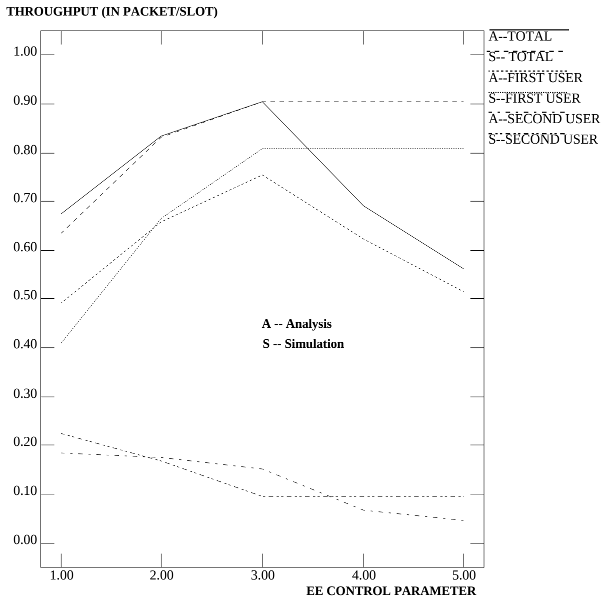

Figure 12: Throughput versus EE ($L = 3$, $\lambda = 1.2$ packet/slot).

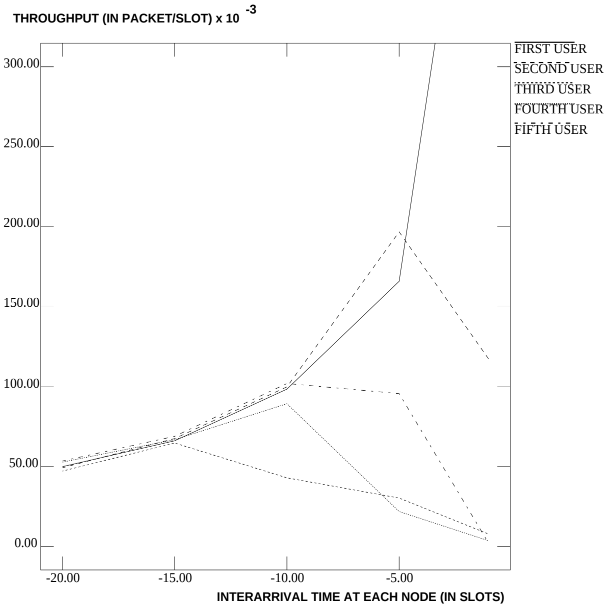

Figure 13: Throughput versus offered load ($Q = 10$, uncontrolled).

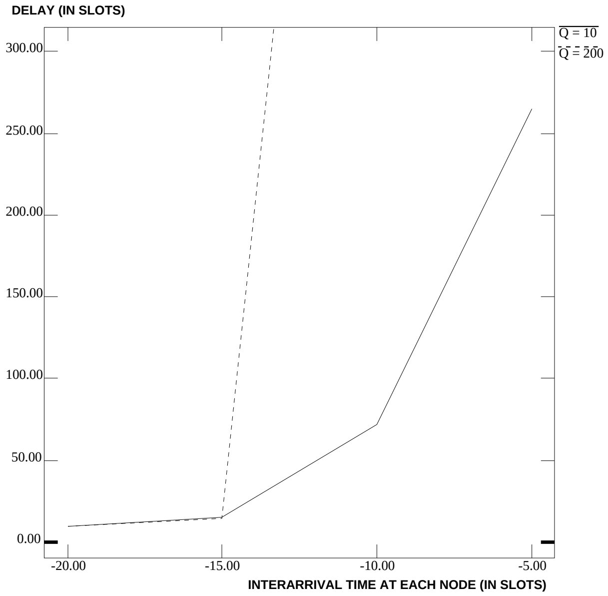

Figure 15: Fifth user delay versus offered load.

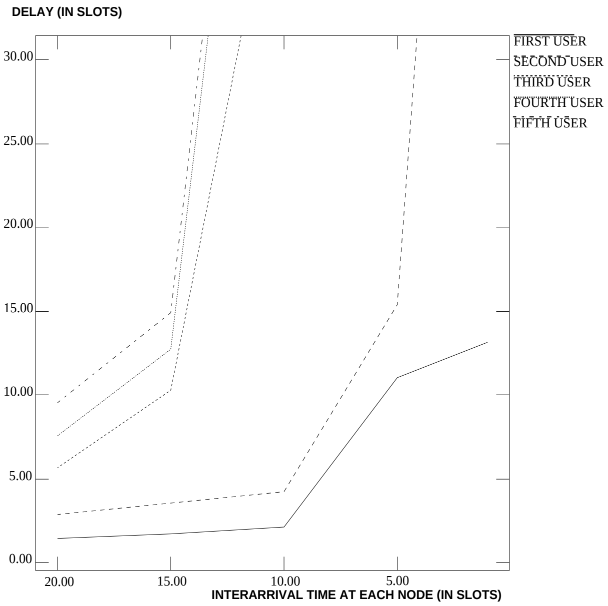

Figure 14: Delays versus offered load ($Q = 10$, uncontrolled).

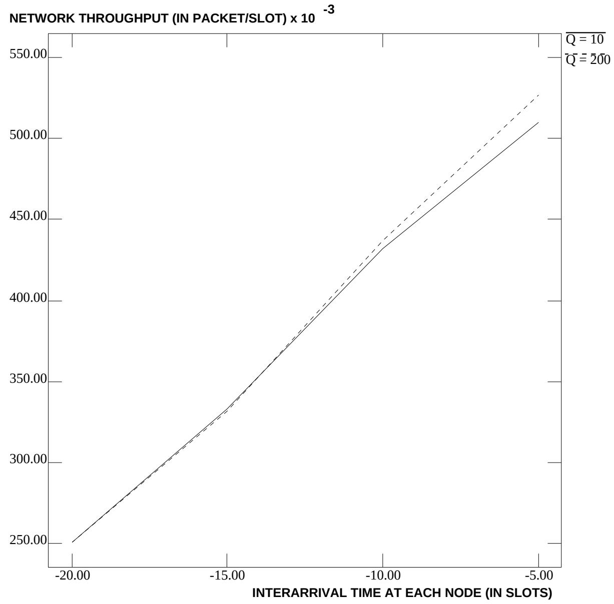

Figure 16: Network throughput versus offered load.

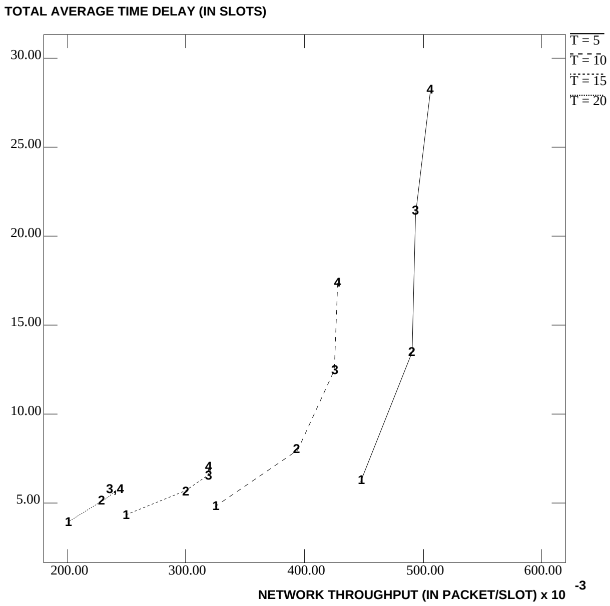

Figure 17: Total average time delay versus network throughput (T denotes average interarrival time at each node (in slots), total offered load = $5/T$ (in packet/slot), $Q = 10$, a point is labeled i = 1, 2, 3, 4 when $EE(i)$ is invoked).

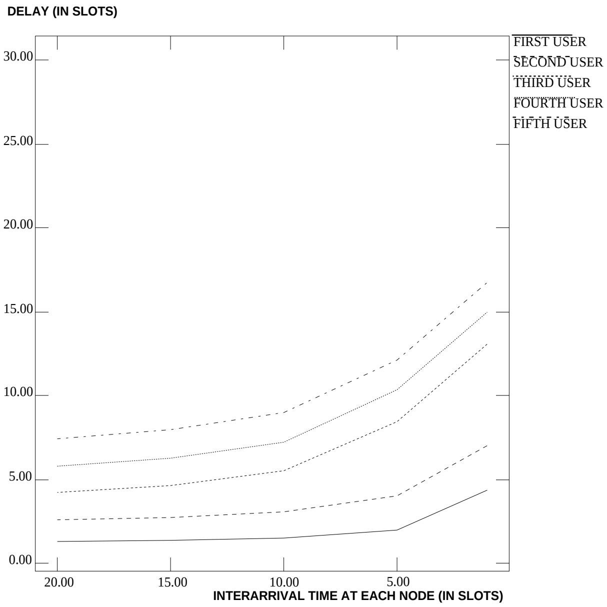

Figure 19: Delays versus offered load ($Q = 10$, with $EE(1)$).

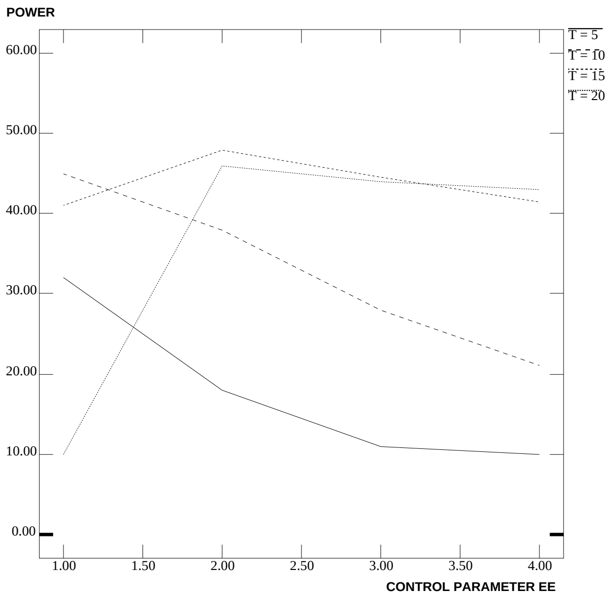

Figure 18: Power versus EE ($Q = 10$).

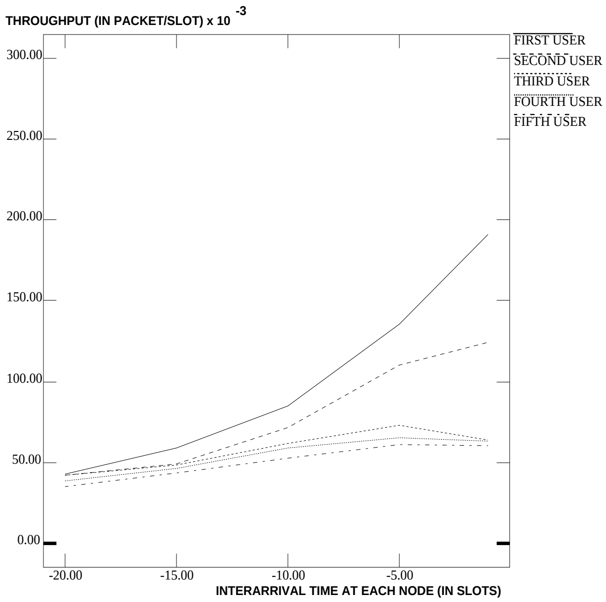

Figure 20: Throughput versus offered load ($Q = 10$, with $EE(1)$).

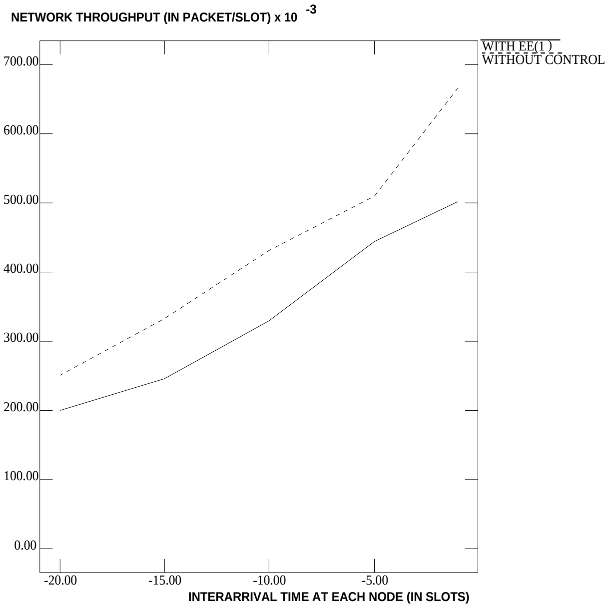

Figure 21: Network throughput versus offered load ($Q = 10$).

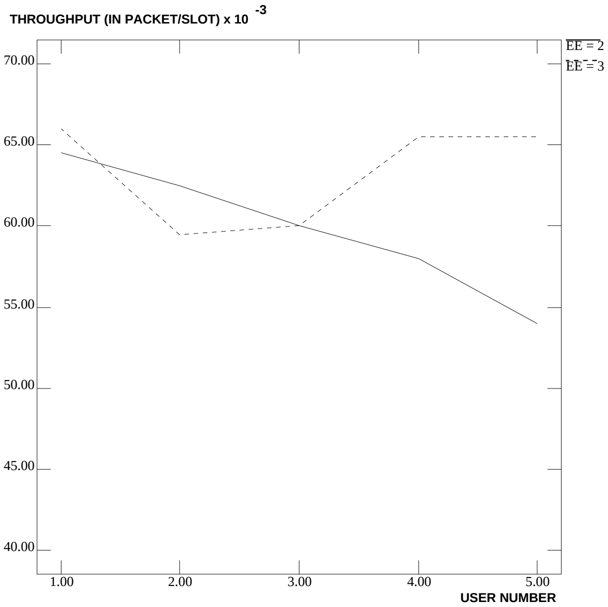

Figure 22: Throughput rates when $EE = 2$ and $EE = 3$ ($Q = 10$, $T = 15$ slots).