Loss Network Models and Multiple Metric Performance Sensitivity Analysis for Mobile Wireless Multi-hop Networks

Invited Paper *

John S. Baras
baras@umd.edu

Senni Perumal
senni.perumal@aims-sys.com

Vahid Tabatabaee
vahidt@umd.edu

Kiran Somasundaram
kirans@umd.edu

Punyaslok Purkayastha
punya@isr.umd.edu

AIMS, Inc. 6213 Executive Blvd., Rockville, MD 20852
and

Institute for Systems Research, University of Maryland College Park, MD 20742

ABSTRACT

We develop and evaluate a new method for estimating and optimizing various performance metrics of mobile wireless multi-hop networks, including MANETs. The method utilizes approximate (throughput) loss model that couples the physical, MAC and routing layers effects. The model provides quantitative statistical relations between the loss parameters that are used to characterize multiuser interference and physical path conditions on the one hand and the traffic rates between origin destination pairs on the other. The model considers effects of the hidden nodes, node scheduling algorithms, MAC and PHY layer failures and unsuccessful packet transmission attempts at the MAC layer in arbitrary time varying network topologies where multiple paths share nodes. The method then applies Automatic Differentiation (AD) to these implicit performance models, to compute sensitivities of various performance metrics with respect to network parameters. We demonstrate the method by applying it to time varying mobile network topologies, including reduced connectivity instances, with both random access MAC (contention mode of the 802.11) as well as reservation based MAC (USAP TDMA based protocol). We analyze throughput, delay and packet loss as metrics and investigate metric optimization and tradeoff analysis. Finally we provide numerical results for realistic mobile networks with time varying topologies.

*This work was supported by SBIR Phase II contract W15P7T-07-C-P613 from U.S. Army CECOM to AIMS, Inc.

WICON'08 , November 17-19, 2008, Maui, Hawaii, USA

Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design—*Wireless Communication, Network Communications, Network Topology*; C.4 [**Performance of Systems**]: Modeling techniques; C.4 [**Performance of Systems**]: Performance attributes

1. INTRODUCTION

Despite recent interest and progress in multi-hop wireless networks, we still lack systematic methodologies and tools that would allow for the efficient design and dimensioning of such networks with the provision of accurate performance bounds. The main reason for this is the different nature of wired and wireless networks rendering the use of wired network techniques inappropriate for the case of wireless networks. Key quantities, such as the link capacity, that remain constant in a wired network, vary in wireless communication environments with the transmission power, the interference, the node mobility and the channel condition.

In this work, we develop hybrid (analytical and numerical) models which can efficiently approximate the performance of a wireless network. These models can assist us to design wireless networks and protocols, and to predict their performance. It is possible to develop packet level simulation tools based on physical (PHY) and medium access control (MAC) layer models using various software packages. However the packet level simulation of multi-hop wireless networks with the appropriate PHY and MAC layer modeling turns out to be too complex and time consuming for the design and analysis of wireless networks in realistic settings. Our objective is to develop low complexity combined analytical and computational (numerical) models, which can efficiently *approximate* the performance of wireless networks. Such models have several applications in the design and analysis of wireless networks

Our approach is based on fixed point methods and loss network models for performance evaluation and optimiza-

Digital Object Identifier: 10.4108/ICST.WICON2008.4970
http://dx.doi.org/10.4108/ICST.WICON2008.4970

tion. Loss network models [8] were originally used to compute blocking probabilities in circuit switched networks [7] and later were extended to model and design ATM networks [6, 13, 2, 9]. In [9] reduced load approximations were used effectively to evaluate quite complex ATM networks, with complex and adaptive routing protocols, and multi-service multi-rate traffic (different service requirements). The main challenge in developing loss network models for wireless networks is the coupling between wireless links. This coupling is due to the transmission interference between different nodes in proximity with each other.

The main mathematical tools we use are fixed point methods and AD for sensitivity analysis. We assume we know the exogenous traffic rate for each source-destination pair, the set of paths and the fraction of traffic traversing each path in the network. In our approach, we model routing, scheduling, PHY and MAC layers of the network architecture. However, in this paper the emphasis is on the MAC layer modeling. We formulate a set of equations modeling the implicit dependencies between the various layers of the protocol stack. With this implicit model we are able to recover network performance parameters such as throughput and delay.

We develop fixed point models for both random access and reservation-based medium access wireless networks. We briefly review our methodology for throughput analysis of 802.11, which is presented in [3], and explain how we can use the same model for delay analysis. For the reservation based MAC protocol, we consider the USAP protocol [14]. The USAP protocol performs both *hard-scheduling*, which is end-to-end reservation of the link capacity for a call on the entire path, and *soft-scheduling*, which is per-hop scheduling of links for single packets after the arrival of the packet to a node.

We use AD (a design tool which provides gradients of software-defined functions), to compute the sensitivities of the network performance parameters. This sensitivity analysis is used to evaluate the resilience and robustness of the solution. The generated *implicit* analytical model, based on the fixed point iterations, is the input to the AD. The AD provides the partial derivatives of the performance metric (e.g. throughput and delay) with respect to defined input parameters (i.e. design variables or parameters). This method allows for very complex design parameters to be implicitly embedded in the input function to the AD module. We further use these sensitivities to compute the optimal load distribution among multiple paths to maximize the network throughput.

The rest of the paper is organized as follows: Section 2 describes the general frame work and briefly review the simple scheduling and routing models that we use in this paper. Section 3 presents the model for 802.11 based MAC models. Section 4 introduces the USAP protocol. Section 5 and 6 describe our fixed point models for the hard-scheduling and soft-scheduling modes of USAP respectively. Section 7 discusses how we use Automatic Differentiation in the current framework for performance metric sensitivity computations. Finally, section 8 provides simulation results and illustrates how the fixed point models can be used for design and performance analysis of wireless networks.

2. GENERAL FRAMEWORK

Figure 1 depicts the main blocks of our performance model

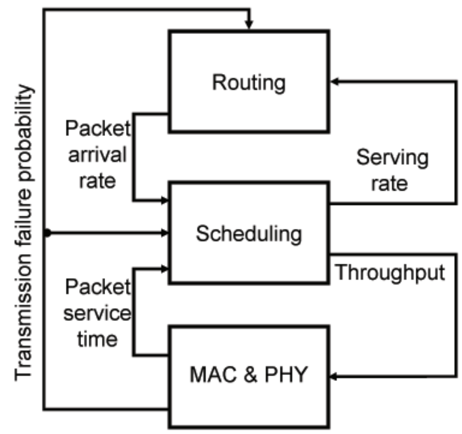

Figure 1: Cross-layer interdependence. All parameters are computed for each node of a path.

for wireless networks and their interdependencies. For each block, a separate model is derived: For the PHY and MAC layer we present a set of equations that model a specific MAC and PHY protocol for a given network topology, which allows us to implicitly express the transmission loss parameters (transmission failure probability) and the average packet service time for each path on each node as a function of the node throughput. The throughput of a path p at a node i is the fraction of time that node i spends in serving path p packets. The routing model derives (computes) the arrival rate for each path at each node as a function of the loss parameters and the serving rates. Finally, the scheduler model computes the serving rate and throughput of path p packets at node i as a function of the packet arrival rates and the loss parameters.

These three sets of equations are coupled iteratively in a fixed point setting, until they converge to a consistent set of solutions. The solution provides an approximation to the packet loss and service time per link and the throughput (outgoing to the incoming traffic ratio) of the network for each connection. In the rest of this section, we briefly review our methodology for modeling the routing and scheduling layers.

The Scheduler Modeling: We consider a network that consists of N nodes and a path set P that is used to forward traffic between the source destination (S-D) pairs in the network. Let P_i denotes the set of the paths that goes through a node i. The scheduler behavior is specified by the scheduler coefficient $k_{i,p}$, which is the average serving rate of path p packets at node i. For simplicity, we assume that all packets have the same length. Let $\lambda_{i,p}$ be the arrival rate and $T_{i,p}$ be the service time of path p packets at node i. *The service time, $T_{i,p}$ is the time that node i scheduler spends serving a path p packet, and starts from the time that the scheduler selects a path p packet to be served and not from the time that the packet becomes head of the queue.* Note that the transmission time includes the retransmissions of a packet too. For instance, in the 802.11 MAC layer, if a transmission fails the packet will be retransmitted up to m times and after that the packet will be discarded.

The scheduling rate is a function of MAC and PHY layer packet failure probabilities. In the 802.11 RTS/CTS pro-

Digital Object Identifier: 10.4108/ICST.WICON2008.4970
http://dx.doi.org/10.4108/ICST.WICON2008.4970

tocol there are two stages for packet transmission: in the first stage is a RTS-CTS transfer and the second is Data packet-ACK transfer . The probability of transmission failure (PHY or MAC layer) for a packet of path p at node i is denoted by $\beta_{i,p}$. The total average throughput $\bar{\rho_i}$, of node i, is,

$$\bar{\rho_i} = \sum_{p \in P_i} k_{i,p} E(T_{i,p}). \tag{1}$$

In order to model a FCFS queueing policy, we assume that the scheduler coefficients are:

$$k_{i,p} = \begin{cases} \lambda_{i,p} & \text{if } \sum\limits_{p' \in P_i} \lambda_{i,p'} E(T_{i,p'}) \leq 1 \\ \frac{\lambda_{i,p}}{\sum\limits_{p' \in P_i} \lambda_{i,p'} E(T_{i,p'})} & \text{otherwise} \end{cases} \tag{2}$$

As described in (2) if utilization of node i is less than one, we can serve all incoming packets. However if the utilization of node i is greater than one, the scheduling rates can be obtained by normalizing the arrival rates by the average utilization to account for the server busy time. Further, the fraction of time $\rho_{i,p}$ that node i is serving path p packets is given by

$$\rho_{i,p} = k_{i,p} E(T_{i,p}). \tag{3}$$

Routing Modeling: In this paper, we consider a simple source-based multi-path routing methodology. In [11], we present our component based methodology for modeling of more complex wireless routing protocols such as OLSR [5]. The routing model specifies a fixed set of paths and the fraction of incoming traffic that is sent over each path at the source node. Due to losses the incoming traffic rate at successive nodes of a path is non-increasing with every hop. Let $h_{i,P}$ be the next hop node of i in path p. The incoming traffic rates of $h_{i,P}$ are derived from the scheduling and loss rates of its upstream link:

$$\lambda_{h_{i,p},p} = k_{i,p}(1 - \beta_{i,p}^m) \text{ for all } i, p. \tag{4}$$

3. RANDOM ACCESS MAC LAYER MODELING

The main output parameters for the random access modeling are $\beta_{i,p}$ and $E(T_{i,p})$. The details of the model and equations are given in [3]. This set of equations will be used as an implicit function to derive loss parameters and packet average service times from the node throughput. We consider the 802.11 MAC layer with RTS/CTS mechanism. This model accounts for the effect of hidden nodes and multiple paths that share nodes in the network. Further, the model accounts for a finite attempt factor m after which the MAC layer packet is discarded.

The packet service time, $T_{i,p}$ is the time to finish a *successful or unsuccessful* transmission of a path p packet at node i, *after* it is scheduled for transmission at node i. The average service time $E(T_{i,p})$ has four components: $d_{i,p}$ is the time spent for successful transmission of path p packets at node i, $u_{i,p}$ is the average time consumed for successful transmission of node i neighbors, $b_{i,p}$ is the average back-off time of node i for path p packets, $c_{i,p}$ is the average time spent in failed transmissions.

$$E(T_{i,p}) = (1 - \beta_{i,p}^m) d_{i,p} + u_{i,p} + b_{i,p} + c_{i,p} \tag{5}$$

Delay Analysis: We start with the total throughput computed in equation (1). If we model the queue at each node as an M/M/1/N queue then the probability of having packets in the queue is,

$$\pi_{i,n} = \frac{(1 - \bar{\rho_i})\bar{\rho_i}^n}{1 - \bar{\rho_i}^{N+1}} \tag{6}$$

The probability of dropping a path p packet due to congestion at node i is, $\gamma_{i,p} = \lambda_{i,p}\pi_{i,N}$. The expected service time of a packet at node i is:

$$S_i = \frac{\sum\limits_{p \in P_i} \lambda_{i,p} E(T_{i,p})}{\sum\limits_{p \in P_i} \lambda_{i,p}} \tag{7}$$

The expected queue length is, $Q_i = \sum\limits_{n=0}^{N} n\pi_{i,n}$. The expected waiting time for a packet in the queue is, $\tau_i = S_i Q_i$. The expected delay for a packet from path p at node i is,

$$D_{i,p} = \tau_i + E(T_{i,p}) \tag{8}$$

We can compute the delay over each path of the network by summing up the corresponding delay of the nodes in the path.

4. RESERVATION BASED MAC

In the previous section, we discussed random access MAC modeling. The main advantage of random access MAC protocols are their simplicity since they need minimal control message exchange between the nodes. Hence, these protocols are very efficient when the system throughput and probability of collision is low. On the other hand, if the traffic rate is high, or when we require predictable and reliable performance the reservation based MAC protocols are more efficient. That is the main reason behind the recent interest in reservation based MAC protocols in both commercial and military applications. In reservation based MAC protocols, the communication channel is usually divided into two separate sections for control and data communication. The control channel is used for coordination between the nodes and reservation of the data channel to avoid collisions. In the following, we will briefly review USAP (Unifying Slot Assignment Protocol) which is a dynamic resource allocation protocol for mobile multihop multichannel wireless networking.

USAP MAC Protocol [14]: Rockwell Collins Corp. has developed a wireless ad-hoc protocol suite for tactical battlefield, which is called Wireless-wideband Networking Engine (WNE). WNE channel access is also adapted to the Mobile Data Link (MDL) layer of the Joint Tactical Radio System (JTRS) Wideband Networking Waveform (WNW). WNE partitions the channel in time and frequency. USAP is the distributed allocation protocol that is used in WNE.

USAP constructs a periodic frame structure as shown in Fig. 2 for communication. The frame length is 125 ms. The bootstrap minislots are pre-allocated to nodes for exchange of information related to network management and are used for reservation of the data channel time-frequency segments which we call cells. Each node informs neighbors about reserved slots and channels using these minislots. Broadcast slots support multicast/broadcast data transmission and we do not consider them in our modeling in this paper. Reservation slots support unicast data traffic.

USAP frame reservation slots consists of $M \times F$ cells, where M is the number of time slots, and F is the number of frequency channels in a frame. Once a cell is assigned

Digital Object Identifier: 10.4108/ICST.WICON2008.4970
http://dx.doi.org/10.4108/ICST.WICON2008.4970

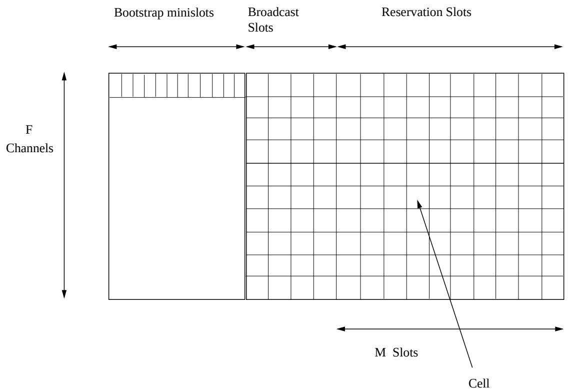

Figure 2: The USAP TDMA Frame Structure

to link (i, j), there is no contention as no node in the two-hop neighborhood transmits simultaneously. Nodes i and j cannot transmit or receive on any other frequency channel corresponding to that time slot.

In the control channel (bootstrap mini-slots) every node broadcasts the cells that are reserved for transmission and reception by itself and its neighbors. In this way, every node acquires information about the reserved slots in its 2-hop neighborhood. Let $T(l)$ and $R(l)$ denote the transmitting and the receiving node on the two ends of link l. To avoid collision, $T(l)$ reservation is based on the following rules:
1- $T(l)$ cannot reserve cells on those time slots which already have scheduled incoming and outgoing cell transmissions to and from $T(l)$ and $R(l)$.
2- $T(l)$ cannot reserve cells used by incoming call transmissions to the neighbors of $T(l)$.
3- $T(l)$ cannot reserve cells used by outgoing call transmissions from the neighbors of $R(l)$.

These rules form the basis for our link capacity approximation that will be described later.

The USAP MAC layer can function under a connection-oriented (hard-scheduling) or connection-less (soft-scheduling) framework. In the hard-scheduling mode cells are reserved for the duration of call on all links of the path from the source to destination. For the soft-scheduling, as the packets reach a node they are queued and reservation is made per-packet for transmission over the link to the next-hop.

The performance metric for the hard-scheduling case is the percentage of calls blocked for each connection. A call is blocked if there is not enough available capacity (cells in frame) on all links of the path. There is no significant queueing for the hard-scheduling case; hence, delay is not an essential performance metric here.

For the soft-scheduling case the performance metric is both delay and throughput. Since there is queueing and limited capacity on the links; we will have packet loss due to congestion and throughput specifies the percentage of the delivered traffic. Qeueing also results in delay that should be studied.

5. USAP HARD-SCHEDULING MODEL

We first briefly recollect the key equations in the loss network approximation for computation of the blocking probability for an incoming virtual circuit connection in a network. We are given the statistics of the ongoing connections (with associated source-destination pairs) in the network, and the routes assigned to such connections.

We assume that calls arrival for route r is a Poisson process with rate ν_r. The holding time for a call is exponential with mean $1/\mu_r$. The demand for a call in terms of the number of (reserved) cells per frame is n_r. The route r calls arrive to link l with offered load

$$\rho_{l,r} = \frac{\nu_r}{\mu_r} \prod_{i \in r/\{l\}} (1 - B_{i,r}). \tag{9}$$

$B_{l,r}$ is the probability of blocking a call on link $l \in r$. The blocking probability, for a connection traversing through route r towards its destination, is given by

$$L_r = 1 - \Pi_{l \in r}(1 - B_{l,r}). \tag{10}$$

Denote $Q_r\left[C; \rho_{r'}, r' \in R_l\right]$ the blocking probability for route r calls on a link l with capacity C and routes set R_l,

$$Q_r\left[C; \rho_{r'}, r' \in R_l\right] = 1 - \sum_{c=0}^{c=C-n_r} q(c) \tag{11}$$

where the $q(c)'s$, are the probabilities of having c cells occupied on link l (chapter 2 of [12]).

In the wireless network case the link capacity is not fixed, and it depends on the number of ongoing connections in the 2-hop neighborhood of the link. For any given value of link capacity m, the occupancy probabilities can be computed using the standard recursive stochastic knapsack algorithm (chapter 2 of [12]). Assume that the link l capacity is between $C_l^{\min}$ and $C_l^{\max}$ with given probability distribution. Hence,

$$B_{l,r} = \sum_{m=C_l^{\min}}^{C_l^{\max}} \left(Pr[C_l = m] Q_r\left[m; \rho_{l,r'}, r' \in R_l\right] \right) \tag{12}$$

From the occupancy probabilities we can compute η_{ij}, the average number of cells that are reserved by each link (i, j):

$$\eta_{ij} = \sum_{C_l^{\min}}^{C_l^{\max}} Pr[C_l = m] \sum_{c=0}^{m} c q_m(c) \tag{13}$$

In order to obtain the blocking probabilities, we need to estimate the link capacity probabilities $P[C_l = m]$.

5.1 Link Capacity Estimation

Define the following quantities for node i: $\Gamma_i^T = \sum_{k \in N(i)} \eta_{ik}$ and $\Gamma_i^R = \sum_{k \in N(i)} \eta_{ki}$, which denote the average number of reserved slots used by i in transmitting to and receiving from its neighbors respectively. Similarly, we define Γ_j^T and Γ_j^R quantities for the node j.

Recall that once a cell (a slot-channel element) is reserved by node i for transmitting or receiving, no other channel can be used for transmitting or receiving by i at that time slot. Let $S_{ij}^i = \Gamma_i^T + \Gamma_i^R - \eta_{ij}$; S_{ij}^i gives the number of slots reserved for transmission to or from neighbor nodes other than node j. Similarly, define the quantity $S_{ij}^j = \Gamma_j^T + \Gamma_j^R - \eta_{ij}$. The minimum and maximum number of cells blocked by i and j communication are,

$$R_{\min}^1(i,j) = \max(S_{ij}^i, S_{ij}^j) \times F, \tag{14}$$

$$\text{and } R_{\max}^1(i,j) = (S_{ij}^i + S_{ij}^j - \eta_{ji}) \times F. \tag{15}$$

Digital Object Identifier: 10.4108/ICST.WICON2008.4970
http://dx.doi.org/10.4108/ICST.WICON2008.4970

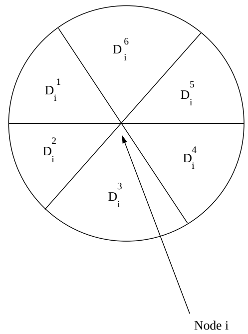

Figure 3: Domains around node i

We now model the effect of communication in the neighborhood of node i. We form regions around node i such that all nodes in that region are in the transmission range of each other. The regions are shown in Figure 3 (the radius of the circle is equal to the transmission range of nodes). The regions are the six sectors of equal area, and are denoted by $D_i^l, l = 1, \ldots, 6$. Let $\Theta^R_{D_i^l} = \sum_{k \in D_i^l, k \neq j} (\Gamma_k^R - \eta_{ik})$; $\Gamma_k^R - \eta_{ik}$ gives the number of slots used by node k for receiving transmissions to it, and excluding the transmissions to it from node i (the term $k = j$ is already taken into account in the computations of $R^1_{\min}(i,j)$ and $R^1_{\max}(i,j)$). Define

$$R^2_{\min}(i,j) = \max_l \Big(\Theta^R_{D_i^l} - \sum_{k \in D_i^l, k \neq j} \eta_{jk} \Big) \tag{16}$$

$$R^2_{\max}(i,j) = \sum_l \Big(\Theta^R_{D_i^l} - \sum_{k \in D_i^l, k \neq j} \eta_{jk} \Big) \tag{17}$$

Similarly, let's define $\Theta^T_{D_j^l} = \sum_{k \in D_j^l, k \neq i} (\Gamma_k^T - \eta_{kj})$; $\Gamma_k^T - \eta_{kj}$ gives the number of slots used by node k for transmission to other nodes excluding node j (the term $k = i$ is already taken into account in the computations of $R^1_{\min}(i,j)$ and $R^1_{\max}(i,j)$). Let us define

$$R^3_{\min}(i,j) = \max_l \Big(\Theta^T_{D_j^l} - \sum_{k \in D_j^l, k \neq i} \eta_{ki} \Big) \tag{18}$$

$$R^3_{\max}(i,j) = \sum_l \Big(\Theta^T_{D_j^l} - \sum_{k \in D_j^l, k \neq i} \eta_{ki} \Big) \tag{19}$$

The lower and upper approximations for C_l are given by

$$C_l^{\min} = M \times F - \Big(R^1_{\max}(i,j) + R^2_{\max}(i,j) + R^3_{\max}(i,j) \Big) \tag{20}$$

$$C_l^{\max} = M \times F - \Big(R^1_{\min}(i,j) + R^2_{\min}(i,j) + R^3_{\min}(i,j) \Big) \tag{21}$$

As a first approximation, we assume that C_l has a uniform probability mass function over the finite set $\{C_l^{\min}, \ldots, C_l^{\max}\}$, and carry out our computations for the blocking probability.

6. USAP SOFT-SCHEDULING MODEL

In the soft-scheduling case, the reservations are done as the packet arrives to a node. Packets will be queued until they are scheduled on a link. If packets arrive to a full buffer they are dropped. The modeling is similar to the 802.11 case, since the data rate decreases due to losses along the path. Losses are due to PHY layer failure and buffer over-flow in the node queues. We are assuming that USAP takes care of the contention, and hence there is no loss due to collisions. We use the same methodology that we discussed in the hard-scheduling case to estimate the link capacity distribution.

Denote the outgoing link of path p at node i with $l_{i,p}$ and the capacity of that link with $C_{l_{i,p}}$. The capacity unit is in cells/sec. The scheduling rates are,

$$k_{i,p} = \begin{cases} \lambda_{i,p} & \text{if } \dfrac{\sum_{p' \in P_i} \lambda_{i,p'}}{C_{l_{i,p}}} \leq 1 \\ \dfrac{\lambda_{i,p}}{\sum_{p' \in P_i} \lambda_{i,p'}} C_{l_{i,p}} & \text{otherwise} \end{cases} \tag{22}$$

This equation is very similar to the scheduling equation (2) that was given for the 802.11 node scheduler. Similarly, the next hop arrival rates will be computed based on the routing model equations given in (4), where $\beta_{i,p}$ represents the PHY layer error probability only.

Capacity of a link depends on the traffic that is going in its second order neighborhood. We use the same methodology that we used for USAP with hard-scheduler to estimate the average capacity available over each link. The capacity computations are based on the parameters η_l, which are the average number of cells reserved for link l,

$$\eta_l = t_f \sum_{p \in P_l} k_{i,p} \tag{23}$$

The unit for $k_{i,p}$ is cells/sec and when it is multiplied by the frame time t_f, it gives the number of reserved cells in a frame. We use the same equations presented in the hard-scheduler case to approximate the link capacities.

Equations (4), (22), (23) and the capacity estimation equations form the basis for the fixed point procedure for the USAP soft-scheduling case that we use to derive the traffic rate over each link and the overall throughput of every connection and the network.

Delay Analysis: For an outgoing link l of node i, let ρ_l (the throughput of link l) be the fraction of time that node i is serving packets forwarded on link l,

$$\rho_l = \frac{\sum_{p \in P_l} \lambda_{i,p}}{C_l} \tag{24}$$

Let T_l be the average service time of path p packets at node i. We assume that when a node reserves a time-channel cell for transmission on a link, it will continue transmitting over that link as long as there are waiting packets. Therefore, if a packet arrives into an empty queue, it should first reserve a cell and then start transmission. On the other hand, if a packet does not arrive to an empty queue, it uses the same cell that is reserved before. In the USAP protocol it takes at least 4 frames to reserve a cell: (1)Node i sends the request. (2) Neighbors of i relay the request from i. (3) 2-hop neighbors confirm (accept) the request. (4) Neighbors of i confirm the reservation.

Let t_f be the period of the USAP frames. We assume that the average number of reserved cells in a frame for link l, if there is any reservation for link l, is $n_l = C_l$

The effective capacity for a link can be smaller than the capacity value computed above, since some cells will be missed while the node is reserving the channel for transmission. Once the reservation is done the average serving time for cells is t_f/n_l. If we model a link as an M/M/1/N queue, the

Digital Object Identifier: 10.4108/ICST.WICON2008.4970
http://dx.doi.org/10.4108/ICST.WICON2008.4970

probability that a link queue is empty is:

$$\pi_0(\rho_l, N) = \frac{1-\rho_l}{1-\rho_l^{N+1}} \tag{25}$$

Then, the average service time of a link is:

$$T_l = 4\pi_0(\rho_l, N)t_f + t_f/n_l \tag{26}$$

and the effective capacity of a link is $C_l = 1/T_l$. Combining these equations we have:

$$T_l = 4\frac{1 - T_l \sum_{p\in P_l} \lambda_{i,p}}{1 - \left(\sum_{p\in P_l} \lambda_{i,p}\right)^{N+1}} t_f + \frac{t_f}{n_l} \tag{27}$$

that can be solved to find T_l. We then use ρ_l in the M/M/1/N equations to derive the expected queue length Q_l. The expected delay for a packet from path p at node i is,

$$D_l = T_l(Q_l + 1) \tag{28}$$

7. SENSITIVITIES & AUTOMATIC DIFFERENTIATION FOR DESIGN

Although the fixed point models for random access and reservation based MAC considered previously can provide the basis for performance analysis of a given network configuration, we need a methodology for network configuration and optimization. We use optimal routing design as an example to illustrate our proposed design methodology. Given a set of paths between source-destination pairs, we use the gradient projection method to find the optimal values for the routing parameters (routing probabilities) to maximize the network throughput. The gradient projection method requires iterative computation of the throughput gradient. If the throughput gradients can be computed analytically after convergence of the fixed point iterations, we can use them. But usually the fixed point method provides a computational scheme that, after convergence, describes the performance metric (i.e., throughput) as an implicit function of the design parameters (i.e., routing parameters). Thus, we do not have analytic expressions of the performance metric evaluations, but instead, we have a program that computes the values of the performance metric, while implicitly providing the dependence of the values on the design parameters. We use Automatic Differentiation (AD) to compute the gradients.

AD is a numerical method to compute the derivatives of a program [4]. Using the fact that a computer program is in fact a sequence of primary operations, automatic differentiation records the relationships between them and using the chain rule, it is able to provide the derivative of a function in a short amount of time. We use ADOL-C (Automatic Differentiation by OverLoading in C++), the source code of which is available at [1] to implement automatic differentiation on our fixed point model. Operator Overloading consists of changing the type of the variables involved in the computation to a proprietary type given by the Automatic Differentiation tool to allow it to compute derivatives based on its linked libraries.

ADOL-C computes the derivatives of real-valued variables and operations that take reals into reals. All the parameters in the 802.11 FPA model are real-valued and hence we were able to successfully use AD to compute the throughput gradients. But AD cannot handle integer-valued variables. In the USAP Hard Scheduling models, the link capacity is an integer representing the number of free slots. Hence it in not possible for AD to compute the derivatives correctly.

But for the reduced load approximation of a multi-service loss network, it is possible to analytically calculate the throughput sensitivities using the implied cost formulation (see section 5.7 of [12]). For the USAP Hard Scheduling models, the total throughput $TH(\mathbf{C_l})$ is defined as the total cell demands that are not blocked and depends on the vector of free capacities $\mathbf{C_l}$ over all the link l , i.e.,

$$TH(\mathbf{C_l}) = \sum_{s\in S}\sum_{r=1}^{k_s} n_s \alpha_{rs} \frac{\nu_s}{\mu_s}(1 - L_r)$$

where S is the set of all source-destination connections, k_s is the total number of routing paths for a connection s, and α_{rs} is the fraction of the calls that are routed over path r for connection s. Hence throughput sensitivities are given by the following equations:

$$\begin{aligned} \frac{\partial}{\partial \alpha_{rs}} TH(\mathbf{C_l}) &= \nu_s(1 - L_r)\Big(\frac{n_s}{\mu_s} - c_s\Big) \\ \text{where, } c_s &= \frac{1}{\mu_s}[TH(\mathbf{C_l}) - TH(\mathbf{C_l} - \mathbf{n_{ls}})] \end{aligned}$$

and $\mathbf{n_{ls}}$ is a vector specifying the call demand requirements of connection s over all the links l in the network.

8. RESULTS

Scenario: The scenario considered is a time varying fast moving network of 30 vehicles that head towards a specified rendezvous point. The scenario duration is for 500 seconds with the vehicles moving at speeds between 22-60 mph. The vehicles start off together, then branch into 3 clusters of 10 nodes each due to obstructions in their path (2 steep hillocks), and finally rejoin (see figure 4). Two Aerial Platforms (APs) are used to maintain communication connectivity when the clusters become disconnected. The number and location of the APs are determined by a fast Deterministic Annealing algorithm [10]. From time 0-30s, the grounds nodes move together and are connected to each other. From time 30-420s, the nodes form 3 clusters with cluster 2 (nodes 10-19) moving in between the two hills while clusters 1 (nodes 0-9) and 3 (nodes 20-29) go around the the hills to the left and right of cluster 2 respectively. The clusters start to lose communication connectivity around 75s, then become disconnected from each other, and finally reconnect around 400s. The APs are brought in to provide communication connectivity between the otherwise disconnected clusters from 75-400s. From 400-500s, the ground nodes move together.

The scenario is specified every 5 seconds (the ground nodes move an average of 100 meters in 5s). At every 5 second interval, the ground node positions, the traffic demands (offered load) & routes between source-destination pairs, and the environment conditions are input to the performance models for random access and reservation based channel access. All ground nodes and APs are assumed to have identical omni-directional radios. The radios have a receiver sensitivity of -95dBsm, a receiver threshold of 10dB, and transmit power of 5W. The environment is modeled as a fading channel with a $1/R^{\alpha}$ power attenuation between nodes.

Digital Object Identifier: 10.4108/ICST.WICON2008.4970
http://dx.doi.org/10.4108/ICST.WICON2008.4970

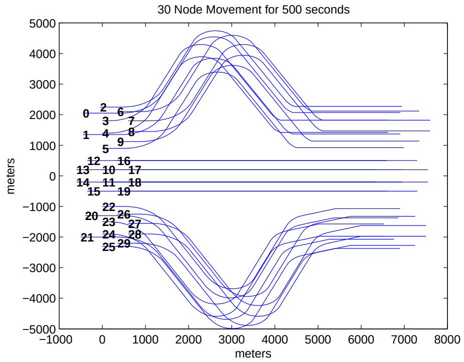

Figure 4: 30 node movement for 500 seconds

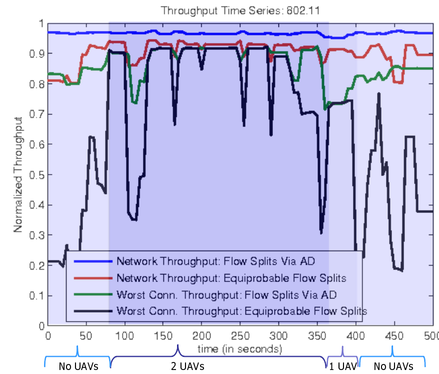

Figure 5: Throughput Time Series: 802.11 model

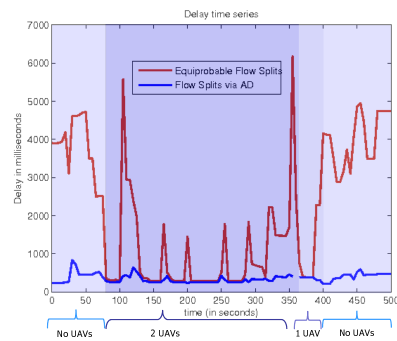

Figure 6: Delay Time Series for Connection 11: 802.11 model

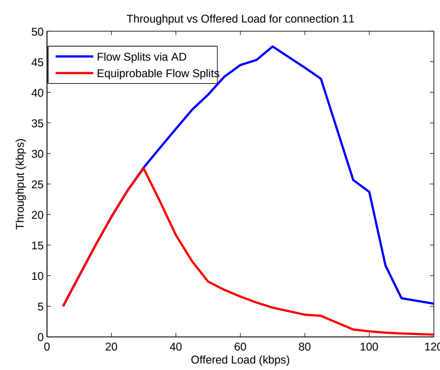

Figure 7: Carried Load vs Offered Load For Connection 11 at snapshot 0: 802.11 model

The radio specification and the path loss exponent α together determine a maximum connectivity distance between the nodes. The path loss exponent α is taken to be 4.5 between ground nodes, 3.9 between ground and aerial nodes, and 3.0 between the aerial nodes. This results in a maximum connectivity distance of 857m between ground nodes, 2423m between ground-aerial nodes, and 25099m between aerial nodes. The maximum channel rate between any two nodes is set to 1 Mbps.

There are 17 source-destination connection pairs chosen in this scenario. The traffic between each source-destination pair is routed via the first K shortest distance paths. There are 13 intra-cluster connections (4 each in cluster 1 and 2; and 5 in cluster 3) each of which have same traffic requirements and with K, the number of paths per connection, equal to 2 or 3. The remaining 4 connections span clusters with K ranging from 2 to 4 paths. Connection 11 is the longest connection and is between source node 20 and destination node 0 (with $K = 4$).

Random Access MAC Layer: The random access 802.11 fixed point model is run on the 30 ground node and 2 AP scenario described above with constant rate traffic between all 17 connections. The 13 intra-cluster connections have source data rate of 100 Kbps. The 4 inter-cluster connections have source data rate between 20 and 100 Kbps. Connection 11, the longest connection, has source data rate of 50 Kbps.

We run the entire scenario first with equiprobable flow splits among the various paths for a connection and then with flow split values optimized using Automatic Differentiation to maximize total throughput. Figure 5 shows the variation of total throughput and worst connection throughput with time for the two cases. Connection 11 with source node 20 and destination node 0 (figure 4) spans all the 3 clusters and exhibits the worst throughput since it has the maximum number of hops per path. Both total throughput and worst connection throughput increase with the flow splits obtained by AD. Figure 6 shows the variation of average delay with time for connection 11 both with equiprobable flow splits and with flow splits determined by AD to maximize throughput. Since lower throughput implies more contention and bigger queues, maximizing the throughput also reduces the average delay.

To capture the effects of offered load on throughput (carried load) and delay, we run the scenario at a particular time (snapshot 0) but with offered loads for all connections scaled by some common factor δ. Figure 7 shows the variation in carried load with offered load for connection 11 at time snapshot 0. We see that as the offered load is increased, the throughput for the longest connection (11) increases to a maximum and then decreases. At low total offered load, the network is operating within its capacity region and hence the throughput increases. But as offered load increases, there is more contention along all the paths and hence throughput decreases. The maximum connection throughput value is both higher and occurs at higher offered load when the flow splits are determined by AD as opposed to equiprobable flow splits. Figure 8 shows the variation in delay with offered

Digital Object Identifier: 10.4108/ICST.WICON2008.4970
http://dx.doi.org/10.4108/ICST.WICON2008.4970

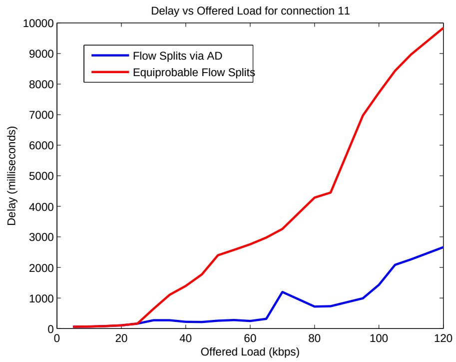

Figure 8: Delay vs Offered Load For Connection 11 at snapshot 0: 802.11 model

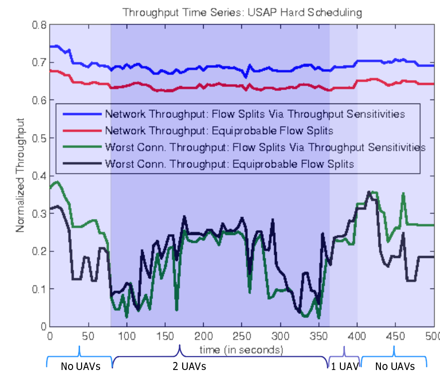

Figure 9: Throughput Time Series: USAP Hard Scheduling

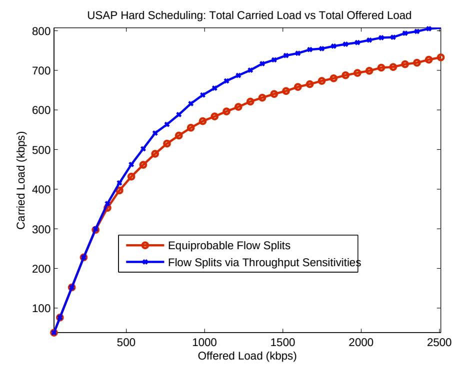

Figure 10: Total Carried Load vs Offered Load: USAP Hard Scheduling

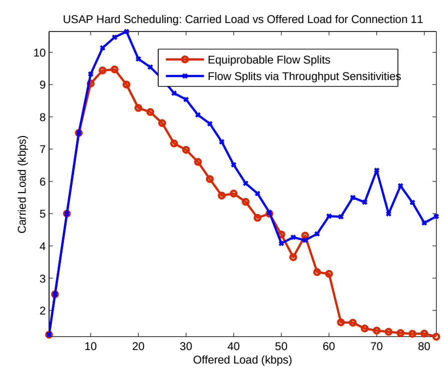

Figure 11: Carried Load vs Offered Load For Connection 11 at snapshot 0: USAP Hard Scheduling

load for connection 11. As the offered load increases, the contention to access the channel in neighborhoods increase, resulting in larger average time to access the channel thereby resulting in higher delay. Moreover as the offered load increases, the queueing delay increases which also contributes to higher delay.

Reservation Based MAC (USAP): The USAP frame period is set to 125ms and the capacity of all the frequency channel is set to 1 Mbps (the same as that used in the 802.11 experiments). The number of frequency channels (F) is set to 2 and the number of reservation time slots (M) is set to 25. Only 50 percent of the USAP frame period is used for reservation slots. Based on the capacity of all channels, M, F, and the fraction of frame period used for reservation slots, 1250 bits can be carried per reservation cell. Hence for a connection to have a call demand (n_r) of 1 reservation slot per frame, the call demand rate (for e.g., the voice coder rate) should be 10 kbps. We assume that the voice coder rate is 10 kbps (hence voice calls use 1 reservation cell per frame) and the voice coder frame period is 125ms.

USAP Hard Scheduling: The USAP Hard Scheduling model is run under the same scenario with all connections comprising of voice calls using 1 reservation cell per frame and holding time of 2 minutes. Since only 50 percent of the USAP frame is available for reservation slots, we decrease the average total call rate per connection (ν_s * hold time * 10 Kbps) to be half that in the 802.11 experiments.

We run the entire scenario first with equiprobable flow splits among the various paths of a connection and then with flow splits determined by the gradient projection using the implied cost formulation for the throughput sensitivities (section 7). Figure 9 shows the variation of total throughput and worst connection throughput (i.e., connection 11) for the two cases. Note the increase in total throughput with the flow splits chosen as per the gradient projection to maximize total throughput.

To find out the effects of offered load on throughput, we ran the scenario at time snapshot 0 but with all connection offered loads scaled by a common factor δ. Figure 10 shows the effect of offered load on total throughput for equiprobable flow splits and flow splits using throughput sensitivities (section 7). The total throughput in both cases saturates to some maximum value as offered load is increased which is the maximum total capacity that the reservation based system can offer. Figure 11 shows the variation of carried load with offered load for connection 11. We see that as the offered load increases, the carried load increases to a maximum and then decreases (with the optimization method doing better that the equiprobable case). Since we are maximizing the total throughput and connection 11 is the longest connection, at high offered loads the shorter connections that use links along its path use more of the link capacity, thereby decreasing the carried load for connection 11.

Digital Object Identifier: 10.4108/ICST.WICON2008.4970
http://dx.doi.org/10.4108/ICST.WICON2008.4970

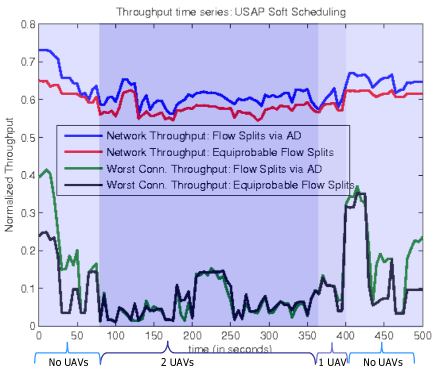

Figure 12: Throughput Time Series: USAP Soft Scheduling

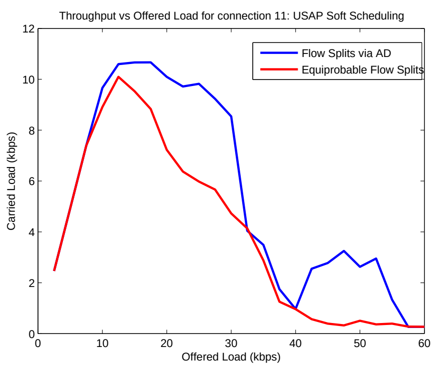

Figure 13: Carried Load vs Offered Load For Connection 11 at snapshot 0: USAP Soft Scheduling

USAP Soft Scheduling: The USAP Soft Scheduling model is also run under the same scenario described with the source arrival rate chosen to be half that of the 802.11 experiments as only 50 percent of the USAP frame is available for reservation slots.

Figure 12 shows the variation of total throughput and connection 11 (worst connection) throughput with time for equiprobable flow splits and flow splits obtained via AD. Note the increase in total throughput with flow splits obtained via AD.

To find out the effects of offered load on throughput and delay, we ran the scenario at a particular time (snapshot 0) but with all connection offered loads scaled by a common factor δ. Figures 13 and 14 show the variation in carried load and delay respectively for connection 11 as the offered load of all connections is increased.

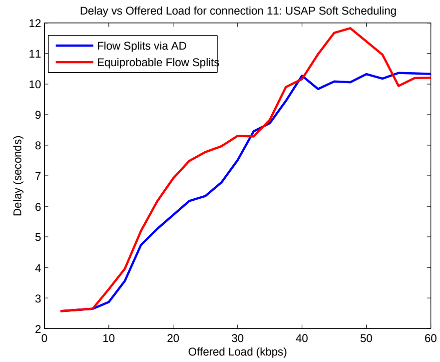

Figure 14: Delay vs Offered Load For Connection 11 at snapshot 0: USAP Soft Scheduling

9. REFERENCES

[1] http://www.math.tu-dresden.de/ adol-c/.

[2] R.Srikanth A.G.Greenberg. Computational techniques for accurate performance evaluation of multirate, multihop communications networks. *IEEE J. Sel. Areas Communications*, 5(2):266–277, Feb 1997.

[3] J. S. Baras, V. Tabatabaee, G. Papageorgiou, and N. Rentz. Modelling and Optimization for Multi-hop Wireless Networks Using Fixed Point and Automatic Differentiation. In *Proceedings of the 6th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt'08), Berlin, Germany*, March 31 - April 4 2008.

[4] Martin Bücker, George Corliss, Paul Hovland, Uwe Naumann, and Boyana Norris. *Automatic Differentiation: Applications, Theory and Implementations*. Birkhäuser, 2006.

[5] T. Clausen, P. Jacquet, A. Laouiti, P. Muhlethaler, and A. Qayyum ans L. Viennot. Optimized Link State Routing Protocol. In *IEEE INMIC Pakistan*, 2001.

[6] J.A.Morrison D.Mitra and K.G.Ramakrishanan. Atm network design and optimization: A multirate loss network framework. *IEEE/ACM Transactions in Networking*, 4(4):531–543, Aug 1996.

[7] F.P.Kelly. Blocking probabilities in large circuit switched networks. *Advances in Applied Probability*, 18(2):473–505, June 1986.

[8] F.P.Kelly. Loss networks. *Annals of Applied Probability*, 1(3):319–378, Aug 1991.

[9] J.S Baras M. Liu. Fixed point approximatio for multirate multihop loss networks with adaptive routing. *IEEE/ACM Trans Networking*, 12(2):361–374, April 2004.

[10] Senni Perumal and J. S. Baras. Aerial Platform Placement Algorithm to Satisfy Connectivity and Capacity Constraints in Wireless Ad-hoc Networks. Submitted to Globecom 2008, Nov 30 - Dec 4 2008.

[11] J.Baras V.Tabatabaee P.Purkayastha and K.Somasundaram. Component based performance modelling of wireless routing protocols. Submitted to ICC 2008, 2008.

[12] Keith W. Ross. *Multiservice Loss Models for Broadband Telecommunication Networks*. Springer Telecommunications Networks and Computer Systems, 1995.

[13] A.Kashper S.Chung and K.W.Ross. Computing approximate blocking probabilities with state-dependent routing. *IEEE/ACM Transactions in Networking*, 1(1):105–115, Feb 1993.

[14] C.D. Young. USAP: a unifying dynamic distributed multichannel TDMA slot assignment protocol. In *Military Communications Conference, 1996. MILCOM '96, Conference Proceedings, IEEE*, Oct. 1996.

Digital Object Identifier: 10.4108/ICST.WICON2008.4970
http://dx.doi.org/10.4108/ICST.WICON2008.4970