

A Time Dependent Performance Model for Multihop Wireless Networks with CBR Traffic

Kunjie Xu, Siriluck Tipmongkonsilp
David Tipper, Prashant Krishnamurthy
Graduate Telecommunications and Networking Program
University of Pittsburgh, Pittsburgh, PA
kux1@pitt.edu, sirilucktip@hotmail.com
{tipper, prashant}@tele.pitt.edu

Yi Qian
Department of Computer
& Electronics Engineering
University of Nebraska-Lincoln, NE
yqian2@unl.edu

Abstract

In this paper, we develop a performance modeling technique for analyzing the time varying network layer queueing behavior of multihop wireless networks with constant bit rate traffic. Our approach is a hybrid of fluid flow queueing modeling and a time varying connectivity matrix. Network queues are modeled using fluid-flow based differential equation models which are solved using numerical methods, while node mobility is modeled using deterministic or stochastic modeling of adjacency matrix elements. Numerical and simulation experiments show that the new approach can provide reasonably accurate results with significant improvements in the computation time compared to standard simulation tools.

1. Introduction

Interest in multihop wireless networks, such as wireless mesh networks (WMN) [1], vehicular ad-hoc networks (VANET) [2], or mobile ad-hoc networks (MANET) [3], has increased significantly in recent years. Multihop wireless networks are expected to become an important part of the communications landscape and may work in a fully autonomous scenario or as an extension to an infrastructure network. In multihop wireless networks, nodes must cooperate with one another to dynamically establish routes using wireless links. Routes may involve multiple hops with each node acting as a router. Since network nodes may move arbitrarily, the network topology is expected to change frequently and unpredictably. Hence, multihop wireless networks require highly adaptive protocols and efficient fail-

ure recovery strategies to deal with the frequent topology changes. Multihop wireless networks also inherit the traditional problems of wireless communications and networking (e.g., asymmetric channels and signal propagation, links that are of poor quality in comparison to wired links, hidden terminal and exposed terminal problems, etc.), which when combined with the other unique aspects, such as mobility and lack of infrastructure make their design and development challenging [1] [2] [3].

Fundamental to the design of efficient multihop wireless networks is the ability to estimate and predict the performance of such networks. Traditionally, multihop wireless network performance has been evaluated using discrete event based computer simulations. Most of the literature adopts a steady state simulation methodology, where for a given network scenario the simulation is conducted over a fixed period of time, with multiple runs from different random number seeds and collected data being averaged over the runs. Additionally, the observations gathered during the transient period in each run are usually eliminated to avoid initialization bias.

A weakness of most of the literature on the performance of multihop wireless networks is that steady state analysis techniques are used even though transient or nonstationary periods will occur in the network, especially after a link or node failure. The importance of this transient behavior after link/node failures has been illustrated in several network technologies including circuit switched networks [4], packet switched data networks [5], packet based signaling networks [6], cellular networks [7], and MANETs [8]. This work, taken together, demonstrates that the dominant factor on network performance after a dynamic event such as a link failure is the transient or nonstationary congestion period. Due to the mobility of nodes and their limited battery life, link and node failures are common in multihop wireless networks. Thus, one would expect that

This research was funded in part by the US Army Research Office under the Multi-University Research Initiative (MURI) grant W911NF-07-1-0318.

transient/nonstationary conditions to occur often and likely dominate the performance behavior [8]. Hence, routing schemes, QoS mechanisms, and congestion control techniques designed and evaluated via steady state analysis may not make optimum use of network resources after a failure or during nonstationary periods.

Simulation studies of time varying behavior for such networks are possible [9], though computationally difficult. To study nonstationary behavior, the measurements of quantities observed over small intervals or at specific points in time are important. Therefore, the time average is not a proper approach and ensemble averages are more appropriate. The idea is to construct ensemble average curves of quantities of interest across a set of statistically identical but distinct independent simulation runs, along with the calculated confidence interval. With many such points collected at different time instants, the behavior of the system can be shown as a function of time. However, to assure the accurate portrayal of the actual system, a large number of runs are required resulting in large amounts of CPU time and scalability issues. In summary, while significant progress has been made towards developing multihop wireless network protocols [1, 2, 3] and developing simulation tools [10, 11, 12, 13] to estimate their performance, relatively little work has appeared on performance models which capture their time varying behavior.

In this paper, we develop a new performance modeling technique that focuses on the dynamics of multihop wireless networks, which includes both time varying and steady state behavior. Network queues are modeled using fluid flow based differential equations, which can be solved by using numerical integration methods. Node connectivity, representing topology changes, are integrated into the model using either deterministic or stochastic based connectivity matrix modeling techniques. Our approach is shown to be more scalable than nonstationary discrete event simulations while allowing the same insight into the interaction of network nodes/protocols.

The rest of the paper is organized as follows. Section II provides the details of our modeling approach. Section III presents numerical results from both our modeling approach and for comparison standard simulation. Section IV studies the computational time complexity of our proposed modeling technique and gives numerical results illustrating the advantages of our method in comparison to simulation. Our conclusions and future works are given in Section V.

2 MODELING DYNAMIC BEHAVIOR

The network topology in multihop wireless networks can change dynamically depending on the link connectivity between each node pair. When nodes in the networks are allowed to move arbitrarily, it will lead to frequent changes

in the topology of the queueing network model. For that reason, we propose a hybrid approach to approximate the network performance using two components: (1) a time varying adjacency matrix model to represent the network topology and (2) a fluid flow model based set of differential equations which model the time dependent queueing behavior of each node. In developing a performance model of the network, we start with modeling a single queue and then generalize to a arbitrary queue in a network. We adopt a fluid-flow based approximation technique to describe the time varying behavior of the queue at each network node, with the help of an approximation concept. Specifically, the *Pointwise Stationary Fluid Flow Approximation (PSFFA)* method [8] [18] [19], which models the average number in the system at a queue by one or more nonlinear differential equations which are derived from a pointwise mapping of the steady-state queueing relationship and can be solved numerically using standard numerical integration techniques.

2.1 Network Topology Modeling

Consider a multihop wireless network consisting of M nodes, the network topology in terms of connectivity at any time t is modeled with a $M \times M$ adjacency matrix denoted as $A(t)$.

$$A(t) = \begin{bmatrix} a_{11}(t) & a_{12}(t) & \dots & a_{1M}(t) \\ a_{21}(t) & a_{22}(t) & \dots & a_{2M}(t) \\ \vdots & \vdots & & \vdots \\ a_{M1}(t) & a_{M2}(t) & \dots & a_{MM}(t) \end{bmatrix}$$

$$\text{where } a_{ij}(t) = \begin{cases} 1 & \text{if node } i \text{ and } j \text{ are directly connected} \\ & \text{at time } t \\ 0 & \text{otherwise} \end{cases}$$

The connectivity between two nodes is determined by their radio range which is a function of a variety of factors, such as the distance between the nodes, antenna radiation pattern, power level, geographic terrain, propagation environment, interference, receiver sensitivity, speed, etc. In general, the range of every node is different and the connectivity among nodes can be asymmetric (i.e., $a_{ij}(t) = 1$ while $a_{ji}(t) = 0$). Given an initial placement of the network nodes, the dynamic network topology due to node mobility is reflected in the adjacency matrix by changes in the $a_{ij}(t)$ values with time. Information about node movement and connectivity can be determined from experimentally gathered trace data, a discrete event simulation of a mobility model (e.g., random waypoint) or stochastic/probabilistic models of mobility effects on link connectivity [16] [17]. In the trace based approach, the data is mined for the link connectivity information versus time. In the simulation approach, a mobility model is used to create the network topology dynamics. Specifically, every node pair is checked

for the possible connectivity change based on their current speeds and directions. Note, that changes in speed, direction and power level are also considered events. The event times are placed in chronological order and as time evolves the pair wise connectivity calculation is repeated for every event time and the matrix is changed accordingly. In this way, the adjacency matrix can reflect the topology change dynamically. Note that no traffic need be generated just the node mobility and range of the nodes, this can be accomplished using existing tools such as the BonnMotion simulator [14]. A computationally simpler approach is to model mobility by directly manipulating the elements of the adjacency matrix according to a planned experiment or a stochastic/probabilistic model (as shown in Section III). Such a probabilistic model can be developed either from the mobility model assumptions and analysis [15] or from fitting a statistical model to data gathered from a test bed or simulation (for example, two-state MMPP [17]).

2.2 Fluid Flow Model Background

Consider a single server first in first out queueing system with a nonstationary arrival process, where $\lambda(t)$ represents the ensemble average arrival rate at time t . The model is developed by focusing on the dynamics of the packet queues at the transmission link. With $x(t)$ defined as the state variable representing the ensemble average number in the system at time t , $\dot{x} = dx/dt$ is the rate of change of the state variable with respect to time. According to the flow conservation principle, the rate of change of the average number in the system equals to the difference between the flow in and the flow out of the system at time t , denoted by $f_{in}(t)$ and $f_{out}(t)$ resulting in:

$$\dot{x}(t) = -f_{out}(t) + f_{in}(t) \quad (1)$$

For infinite waiting space queues, the flow in equals the arrival rate $f_{in}(t) = \lambda(t)$. The flow out can be related to the ensemble average utilization of the server as $f_{out}(t) = \mu C \rho(t)$, where $1/\mu$ refers to average packet length, C defines the server capacity and $\rho(t)$ is the server utilization. Thus μC represents the average service rate in number of packets per unit time. This equation can model a wide range of queueing systems as shown in [5] and [19]. The fluid flow equation can then be written in terms of the average arrival rate and the departure rate as:

$$\dot{x}(t) = -\mu C \rho(t) + \lambda(t) \quad (2)$$

The utilization $\rho(t)$ in (2) depends on stochastic modeling assumptions of the queue under study such as traffic arrival process and service time distribution. In general, the exact expression of $\rho(t)$ is difficult to determine. In [19] we proposed the function be determined by an approximation approach - matching the equilibrium point in the

differential equation at particular time instant to the corresponding steady state queueing theory result. To adopt this mapping approach, we assume that the server utilization function can be approximated by the non-negative function $G(x(t))$, which represent the ensemble average utilization of the server at time t as a function of the state variable $x(t)$. Then (2) can be written as:

$$\dot{x}(t) = -\mu C(G(x(t))) + \lambda(t) \quad (3)$$

In general, the function $G(x(t))$ is nonlinear and we can solve for $x(t)$ in (3) using numerical methods as follows. Given an initial condition of the state variable at time zero as $x(0)$ and an approximation of the arrival rate as a constant λ over a small time step $[0, \Delta t]$ (i.e., $\lambda = \lambda(\Delta t/2)$), equation (3) can be solved numerically for state variable at the end of the time interval $x(\Delta t)$, which then becomes an initial condition for the next time step $[\Delta t, 2\Delta t]$. The arrival rate for the new time step is adjusted and the procedure is repeated for each time interval along the time horizon. Numerical studies in [18] have shown that results from the PSFFA model are reasonably accurate. Here we derive the PSFFA model for the queueing system of interest beginning with the general multiple traffic stream case.

2.3 Multiple Traffic Stream Fluid Flow Model

Consider the case of a single queue with S different traffic streams, each with the average arrival rate of $\lambda_1(t), \lambda_2(t), \dots, \lambda_S(t)$ respectively as illustrated in Fig. 1.

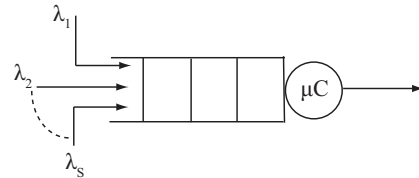


Figure 1. Queueing model with S traffic streams

The aggregate traffic can be considered as one arrival process $\lambda_T(t) = \sum_{l=1}^S \lambda_l(t)$. Let $x_l(t)$ represents the ensemble average number of stream l in the system at time t , the total average number in the system is defined as $x_T(t) = \sum_{l=1}^S x_l(t)$, and the state model in (3) becomes:

$$\dot{x}_T(t) = -\mu C(G(x_T(t))) + \lambda_T(t) \quad (4)$$

We note that the flow conservation principle also applies to each traffic stream. Therefore, a state model can also be developed for each stream with $G(x_l(t), x_T(t))$, as the average utilization function of the link by stream l traffic, as

a function of the total average number in the system x_T and the average number of stream l packets in the system x_l .

$$\dot{x}_l(t) = -\mu C(G_l(x_l(t), x_T(t))) + \lambda_l(t) \quad \forall l = 1, 2, \dots, S \quad (5)$$

Thus the transmission link can be described by a set of S coupled state equations, each representing the traffic behavior of its own stream. Next we tailor this model to the case of periodic traffic streams.

2.4 Modeling CBR Traffic

We focus on modeling a queue with constant bit rate input traffic streams and consider two cases: (1) where all CBR streams have the same data rate and (2) when the CBR streams have different data rates. The two cases are studied in turn below.

Case I: Identical Sources

Consider a queuing model as in Fig. 1 where all the traffic sources have the same period (although not necessarily in synchronization) and the packets have the same size. This model is often referred to as the $S * D/D/1$ queue. The period of the S input streams is measured in server time units D , and the server utilization equals to $\rho = S/D$ which must have $\rho < 1$ for stability.

Following [23] and [25], we model lack of synchronization among the streams by assuming that the first arrival of each flow is independently and uniformly distributed over the first arrival period interval. Let L_t denote the number of customers present at time t and $Q_t(r) = \Pr\{L_t > r\}$ is the survival function or complementary distribution of the number in the system at arbitrary instant t . Further let $A(u, t)$ denote the number of arrivals in an interval (u, t) . As noted in [23] the survival function can be written as:

$$Q_t(r) = \sum_{s=1}^{\infty} p_s(r) \pi_0(r, s) \quad (6)$$

where $p_s(r) = \Pr\{A(t-s, t) = r+s\}$ and $\pi_0(r, s) = \Pr\{\text{system empty at } t-s \mid r+s \text{ arrivals in } (t-s, t)\}$. Noting that the binomial distribution provides the probability of arrivals during an interval. Then $Q_t(r)$ [25] can be written as:

$$Q_t(r) = \sum_{s=1}^{S-r} \left[\binom{S}{r+s} \left(\frac{s}{D}\right)^{r+s} \left(1 - \frac{s}{D}\right)^{S-r-s} \left(\frac{D-S+r}{D-s}\right) \right] \quad \text{for } 0 \leq r < S \quad (7)$$

where the first three terms in the sum represent the number of arrivals and the rightmost term represents the probability the system is initially empty given $r+s$ arrivals. The total average number in the system x_T can be found using

$Q_t(r)$ [23], since the mean of nonnegative variable x_T can be computed as $x_T = \sum_{r=0}^{S-1} Q_t(r)$ [24]. Therefore, for the $S * D/D/1$ queue, x_T is given by:

$$x_T = \sum_{r=0}^{S-1} \sum_{s=1}^{S-r} \left[\binom{S}{r+s} \left(\frac{s}{D}\right)^{r+s} \left(1 - \frac{s}{D}\right)^{S-r-s} \left(\frac{D-S+r}{D-s}\right) \right] \quad \text{for } 0 \leq r < S \quad (8)$$

The above formula can be used to numerically determine x_T for a given traffic mix (i.e., S, D, ρ). We apply a curve fitting approach using the (ρ, x_T) data from (8) to find the utilization function $G(\cdot)$. The resulting $G(x_T(t))$ is in the form of a polynomial (i.e., $G(x_T(t)) = ax_T^n + bx_T^{n-1} + \dots + k$) and can be substituted back into the general fluid flow model (4).

Determining the state model for each traffic stream l at a queue, following the approach of steady state equilibrium matching with $\dot{x}_T(t) = 0$, from (4) and (5), we find $\lambda_T(t) = \mu C G(x_T(t))$ and $\lambda_l(t) = \mu C G_l(x_l(t), x_T(t)) \quad \forall l = 1, 2, \dots, S$, respectively. Solving these equations along with $G(x_T(t))$ determined from data fitting, the average utilization for stream l packets, $G_l(x_l(t), x_T(t))$ can be determined as

$$\begin{aligned} G(x_l(t), x_T(t)) &= \frac{x_l}{x_T} * G(x_T(t)) \\ &= \frac{x_l}{x_T} [ax_T^n + bx_T^{n-1} + \dots + k] \end{aligned} \quad (9)$$

This can be substituted into (5) to provide the multiple traffic stream fluid flow model.

Case II: Non-identical Sources

Now we consider the case where a group of sources with different rates are multiplexed on a transmission link where the total bit rate is less than the transmission capacity to ensure stability. Suppose there are m types of sources with N_i of type i generating traffic with the arrival period of D_i time units, for $i = 1, 2, \dots, m$. This type of queue may be referred to as a $N_1 D_1 + N_2 D_2 + \dots + N_m D_m / D / 1$ queue. Similar to Case I, we rely on numerical analysis of the queue length distribution to find the utilization function $G(\cdot)$. Following the same approach as before, we focus on determining $Q_t(r)$ using (6).

Determining $p_s(r)$ in (6), we note for an interval of length s , a type i stream will generate $[s/D_i]$ arrivals and possibly one additional arrival with probability $\alpha_i = s/D_i - [s/D_i]$. Considering all N_i of type i sources, the total number of arrivals equals to $N_i[s/D_i] + k_i$ where k_i is a random variable representing additional arrivals for each traffic type with the following distribution:

$$\begin{aligned} b_{si}(k) &= \Pr\{k_i = k\} \\ &= \begin{cases} \binom{N_i}{k} \alpha_i^k (1 - \alpha_i)^{N_i - k} & \text{for } 0 \leq k \leq N_i \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (10)$$

Let $q_s(k)$ be defined as the distribution of $\sum k_i$ for all m types, it is given by:

$$q_s(k) = \sum_{\sum_i k_i = k} \prod_{i=1}^m b_{si}(k_i) \quad \text{for } 0 \leq k \leq \sum N_i \quad (11)$$

Then the probability $p_s(r)$ of $r + s$ arrivals during time interval $(t - s, t)$ can be determined from $q_s(k)$ as:

$$p_s(k) = \begin{cases} q_s(r + s - \sum_i N_i[s/D_i]) & \text{for } \sum_i N_i[s/D_i] \leq r + s \leq \sum_i N_i([s/D_i] + 1) \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

The conditional probability of queue being idle at the beginning of the interval $\pi_0(r, s)$ in (6) is difficult to derive. However, bounds for probabilities $Q_t(r)$ can be obtained from the fact that $(1 - \rho) \leq \pi_0 \leq 1$, where ρ corresponds to the server utilization $\sum N_i/D_i$. Substituting in (6), we have the bounds:

$$(1 - \rho) \sum_{s=1}^{\infty} p_s(r) \leq Q_t(r) \leq \sum_{s=1}^{\infty} p_s(r) \quad (13)$$

A tighter upper bound has been derived in [26] as:

$$Q_t(r) \leq \sum_{s=1}^{\infty} \left\{ \sum_{k_i=r+s-d_s} \prod_{i=1}^m \left[b_{si}(k_i) \left(1 - \sum_{i=1}^m \frac{N_i - k_i}{D_i(1 - \alpha_{si})} \right)^+ \right] \right\} \quad (14)$$

with $(x)^+ = \max\{0, x\}$ and $d_s = \sum N_i[s/D_i]$. Noting that $x_T = \sum_{r=0}^{S-1} Q_t(r)$, one can use function $Q_t(r)$ from (13) or (14) to find the corresponding bounds for x_T . After that, curve fitting the data pair (ρ, x_T) results in the polynomial utilization function $G(x_T(t))$, which can be substituted back into the general fluid flow model (5). The result is two fluid flow models, one providing a lower bound and the other an upper bound. As an alternative, one can simplify the computation by assuming homogeneous traffic. For the lower bound, traffic from all sources are fixed with period of $D_{max} = \max\{D_i\}$ and for the upper bound the traffic period equals to $D_{min} = \min\{D_i\}$. We can apply D_{max} and D_{min} into (8) from $S * D/D/1$ to obtain the upper and lower bounds of the data pair (ρ, x_T) and then find out the utilization function for each bound by curve fitting. Once the bounding fluid models for the total traffic flow x_T are found, one can similarly find upper and lower bounding fluid models for each traffic stream using (9).

2.5 Networks with Periodic Traffic

Consider a multihop wireless network consisting of M nodes with all traffic being constant bit rate. An arbitrary

node i is shown in Fig. 2. At each node, there are $M - 1$ possible traffic types, sorted into $M - 1$ classes based on their destinations. We assume that packets are generated at the node i destined for node j according to a deterministic process, (which can be nonstationary), with mean rate $\gamma_i^j(t)$. Let $x_i^j(t)$ denote the average number of packets in the queueing system at node i destined for node j at time t . We denote the average packet length with $1/\mu$ and let C_i denote the transmission capacity of node i . When considering the network as a whole, (5) must be modified to clearly identify the source node i and destination node j for each state variable $x_i^j(t)$, as well as to model traffic being routed through intermediate nodes when a direct link is not accessible. We define $a_{ij}(t)$ in the adjacency matrix to represent node connectivity, as described in section 2.1. Similarly, in order to model multihop routing, we define the routing variable $r_{ik}^j(t)$ as a zero/one indicator variable determined by the routing algorithm, with $r_{ik}^j(t) = 1$ if traffic from node i destined to node j is routed through network node k at time t and $r_{ik}^j(t) = 0$ otherwise.

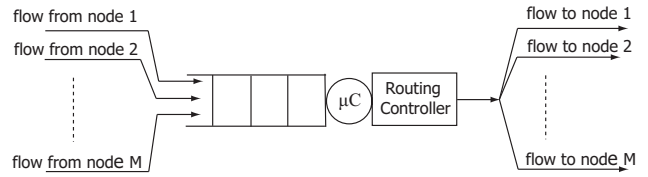


Figure 2. An arbitrary node i queueing model.

Note that, the flow out of node i to node k of a particular traffic class j will depend upon the existence of a direct link between i and k and the routing variables for traffic class j . Hence one must modify the flow out term in (3) to incorporate $a_{ik}(t)$ and $r_{ik}^j(t)$. Specifically, the flow out of node i of class j traffic to node k is given by

$$\begin{aligned} & \text{class } j \text{ traffic flow out of node } i \text{ to node } k \\ &= \mu C_i (G_i^j(x_i^j(t), x_T(t))) (a_{ik}(t) r_{ik}^j(t)) \end{aligned} \quad (15)$$

The flow of class j traffic into the node i queue will consist of traffic generated at node i with rate $\gamma_i^j(t)$ and the flow of class j traffic to node i from other network nodes. For example, the flow of class j traffic into node i from node l is given by

$$\begin{aligned} & \text{class } j \text{ traffic flow into node } i \text{ from node } l \\ &= \mu C_l (G_l^j(x_l^j(t), x_T(t))) (a_{li}(t) r_{li}^j(t)) \end{aligned} \quad (16)$$

In interconnecting queues, the literature [20]-[22] indicates that the output from a queueing system with deterministic service time should be treated as a delayed input to the next stage. This idea is applicable to our model, where the input to the next stage is a superposition of the delayed

input streams from the nearby nodes plus any external arriving traffic. We illustrate the concept by considering a simplified two stage tandem queue model as in Fig. 3(a)-(b), let $x_i(t)$, $\lambda_i(t)$ and $G_i(t)$ be the average number in the system, total average arrival rate and average utilization at node i , respectively. Then, $\lambda_1(t) = \gamma_1(t)$ is the arrival rate to the first queue, and $\mu C G_1(t)$ is the departure rate from the first queue. The departure rate then becomes the input to the second queue with a deterministic propagation delay of service period D_s time units, that is $\lambda_2(t) = \mu C G_1(t - D_s) + \gamma_2(t)$. We can then write a set of fluid flow equations at node 1 and node 2, for Fig. 3(a)-(b) as:

$$\begin{aligned}\dot{x}_1(t) &= -\mu C G_1(t) + \gamma_1(t) \\ \dot{x}_2(t) &= -\mu C G_2(t) + \gamma_2(t) + \mu C G_1(t - D_s)\end{aligned}\quad (17)$$

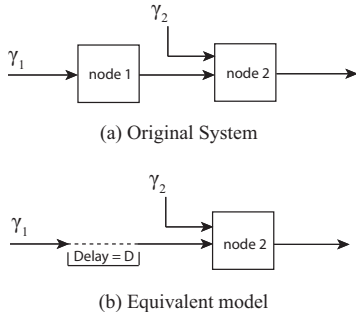


Figure 3. A two-node deterministic service system with its equivalent model.

The general M nodes network hybrid model combining fluid-flow model with routing and connectivity is obtained by summing the flow in and out over all possible nodes:

$$\begin{aligned}\dot{x}_i^j(t) &= -\mu C_i(G_i^j(x_i^j(t), x_T(t)) \sum_{k=1, k \neq i}^M a_{ik}(t) r_{ik}^j(t) + \\ &\gamma_i^j(t) + \sum_{l=1, l \neq i, j}^M \left(\mu C_l(G_l^j(x_l^j(t - D_s), x_T(t - D_s))) \right. \\ &\left. (a_{lk}(t) r_{li}^j(t)) \right) \quad \forall i, j = 1, 2, \dots, M\end{aligned}\quad (18)$$

For a queue with the superposition of periodic arrival streams, the server utilization function $G(\cdot)$ can be written in the form of polynomial expression given in (9). Therefore, the final network fluid flow model is

$$\begin{aligned}\dot{x}_i^j(t) &= -\mu C_i \frac{x_i^j(t)}{x_T(t)} [a x_T^n(t) + b x_T^{n-1}(t) + \dots + k] \\ &\sum_{k=1, k \neq i}^M a_{ik}(t) r_{ik}^j(t) + \gamma_i^j(t) + \sum_{l=1, l \neq i, j}^M \mu C_l \frac{x_l^j(t - D_s)}{x_T(t - D_s)} \\ &[a x_T^n(t - D_s) + b x_T^{n-1}(t - D_s) + \dots + k] (a_{lk}(t) r_{li}^j(t)) \\ &\quad \forall i, j = 1, 2, \dots, M\end{aligned}\quad (19)$$

Given a routing algorithm, connectivity model and traffic information, this model can be solved numerically for the mean queue lengths as a function of time using any standard numerical integration technique.

2.6 Additional Performance Metrics

In addition to the mean queue length versus time, other performance metrics can be evaluated from the model. Here for the sake of brevity we discuss estimation of the end-to-end delay only. Typically, a packet is forwarded from the source via a path which may include several intermediate nodes until it reaches the destination. As a result, the end-to-end delay is the sum of the delays experienced at each node along the way. The delay at a node consists of the processing time, the queuing delay, the transmission time on the link and the propagation time over the link to the next node. In general, the queuing and transmission delay are the main factors.

From Little's theorem, the average number in the system is equivalent to the product of the average arrival rate and the average time in the system. If x denotes the average number in the system, λ is the average arrival rate and W is the average waiting time, then $x = \lambda W$. With the assumption of constant arrival rate over a small step, the change in average waiting time can be related to the rate of change in average number in the system $\dot{W} = \dot{x}/\lambda$. Now consider a path P of $j-1$ hops from source node 1 to destination node j (class j traffic), given by $(1, 2), (2, 3) \dots (j-1, j)$, where $(i, i+1)$ represents a link on the path, for $\forall i = 1, 2, \dots, j-1$. The average node delay at node i for class j traffic on link $(i, i+1)$ is denoted by $W_i^j(t)$ and the total latency of path P can be written as $W_P(t) = \sum_{i=1}^{j-1} W_i^j(t)$. However, as we are interested in the time dependent behavior, the rate of change of this path delay is given by:

$$\dot{W}_P(t) = \sum_{i=1}^{j-1} \dot{W}_i^j = \sum_{i=1}^{j-1} \frac{\dot{x}_i^j(t)}{\lambda_i^j(t)} \quad (20)$$

according to (18), $\lambda_i^j(t)$ can be calculated by

$$\begin{aligned}\lambda_i^j(t) &= \gamma_i^j(t) + \sum_{l=1, l \neq i, j}^M \left(\mu C_l(G_l^j(x_l^j(t - D_s), \right. \\ &\left. x_T(t - D_s))) (a_{lk}(t) r_{li}^j(t)) \right)\end{aligned}\quad (21)$$

3 NUMERICAL RESULTS

We solved the fluid flow based model, presented in the previous section, using the fifth order *Runge-Kutta* numerical integration routine with variable time step Δt in Matlab. For comparison purposes, an equivalent discrete event

simulation is built using OPNET 14.5 [11]. In our OPNET model, each queue of the node is configured as a first in first out (FIFO) queue with infinite buffer size and each traffic stream is buffered at a different subqueue. In addition, we use minimum hop routing for both the fluid flow model and the discrete event simulation. The discrete event simulation results are the average of 4096 runs using the nonstationary simulation methodology of [9] discussed in the introduction.

A simple simulation scenario of three nodes with pre-determined connectivity change between nodes at each time interval during the length of simulation time as illustrated in Fig. 4(a) - (f), is studied first. This topology is used to evaluate the accuracy of our proposed model. In this setup, when the direct link is no longer available, traffic must be rerouted through relay nodes and uses some available portion of the shared link capacity. We set the capacity for all nodes $C_i = 10^4$ bps with packet length $1/\mu = 1250$ bytes, so that the average service rate $1/D_s$ is normalized to one packet per second.

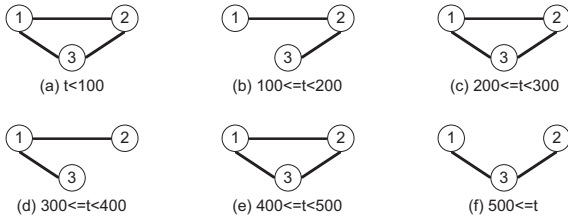


Figure 4. Three node network connectivity scenario.

For case I, the rate of externally arriving packets of each node pair are set to $\gamma_1^2 = \gamma_1^3 = \gamma_2^1 = \gamma_2^3 = \gamma_3^1 = \gamma_3^2 = 0.2$ packets per second but are not synchronized (i.e., the first arrival from each stream is determined from a uniform [0,5] random variable). As described in Section II.C case I, we first compute the data pair (x_T, ρ) from (8). Curve fitting the data pair (ρ, x_T) , the server utilization function is determined as $G(x_T) = 0.0832x_T^3 - 0.4353x_T^2 + 1.0843x_T$. This is then used in (19) to model the network. Fig. 5 shows the results of the effect of topology change on the average number in system and end-to-end delay for traffic at node 1 destined for node 2. For the time interval $t < 100$ sec, all nodes are directly connected. Nodes go through an initial transient period and reaches steady state. For time $100 \leq t < 200$ sec, the link between node 1 and 3 breaks and traffic going through this link has to go through the relay node 2. But traffic x_1^2 is not affected. For the time interval $300 \leq t < 400$ sec, the link between node 2 and node 3 breaks, leading to traffic re-routing and an increase in the number of x_1^2 packets in the system. Due to higher server utilization of node 1, the average queuing time of

each packet at node 1 increases and we can see the rise in end-to-end delay of the traffic from node 1 destined for node 2. During the time $t \geq 500$ sec, link 1-2 breaks and x_1^2 has to go through the relay node 3 to reach the destination. Hence, the end-to-end delay consists of propagation delay in link 1-3 and link 3-2 as well as the queueing delay at node 1 and node 3. The behavior of the other network nodes and traffic streams are similar and not shown here for the sake of brevity.

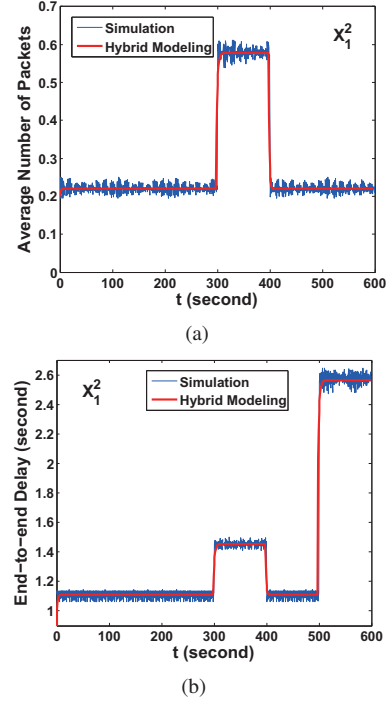


Figure 5. Average number of packets and end-to-end delay of x_1^2 with $S * D/D/1$ queue.

Next we tested the performance of the fluid flow model with case II, the rates of externally arrivals are set as: $\gamma_1^2 = 0.16$ $\gamma_1^3 = 0.2$ $\gamma_2^1 = 0.16$ $\gamma_2^3 = 0.2$ $\gamma_3^1 = 0.16$ $\gamma_3^2 = 0.2$ packets per second. The baseline case when all links are working results in each node being a $N_1 D_1 + N_2 D_2 / D / 1$ queue with $D_1 = 1/0.16 = 6.25$ and $D_2 = 1/0.2 = 5$. Following the procedure described in section II.D case II, we calculate the data pair (x_T, ρ) for the lower bound from the left hand term in (13) and the upper bound from (14). The resulting lower and upper bounds utilization functions are $G_{lower}(x_T) = -63.5754x_T^3 + 49.7702x_T^2 - 7.8764x_T$ and $G_{upper}(x_T) = 0.2061x_T^3 - 0.6099x_T^2 + 1.1257x_T$. The utilization functions are then used in (19) to form two models of the network. Fig. 6 shows the behavior of the traffic at node 2 destined for node 3 (i.e., x_2^3). Here, while all nodes are connected through direct link, the bounds are tight, but when the direct link breaks, a wider gap occurs between bounds. The first reason is that the server utilization ρ of the

relay node becomes higher so that the lower bound generated by (13) drops off. Secondly, (13) does not make use of any information of the k_i combination to tighten the bound. Another way to possibly reduce the gap between bounds is by assuming homogeneous traffic, as discussed in section 2.4 case II and as implemented in the next scenario.

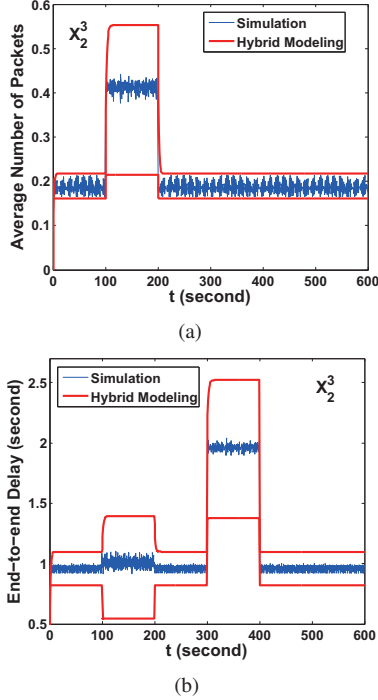


Figure 6. Average number of packets and end-to-end delay of traffic destined for node 3 at node 2

Next we consider a five node network with the random waypoint mobility (RWM) model. Recently in [16] and [17], it is shown that the RWM model link connectivity can be modeled as a two-state Markov process with on-off (connected-disconnected) transition, with both link on and off durations following exponential distributions. This link connectivity model can reproduce the average link stability statistics of RWM model without requiring a detailed node mobility simulation. In this experiment, each link is exponentially distributed on duration with a mean of 50 seconds and off duration with a mean of 10 seconds. The link capacity and packet length remain the same as ones in the three node network scenario. The externally arrival rates are assigned as: $\gamma_1^3 = 0.25$, $\gamma_1^5 = 0.16$, $\gamma_2^5 = 0.18$, $\gamma_3^5 = 0.22$, $\gamma_4^5 = 0.25$. In order to illustrate the typical time varying behavior of the network, we randomly select the time interval [2100 2200] from a 6000 seconds of the total experiment duration. The link connectivity as determined by the RWM for the time interval is shown in Fig. 7. In the following discussion, we focus on the traffic from node 1

destined for node 5 and all the routes of class 5 traffic going through node 1 are marked by dotted lines in the figure.

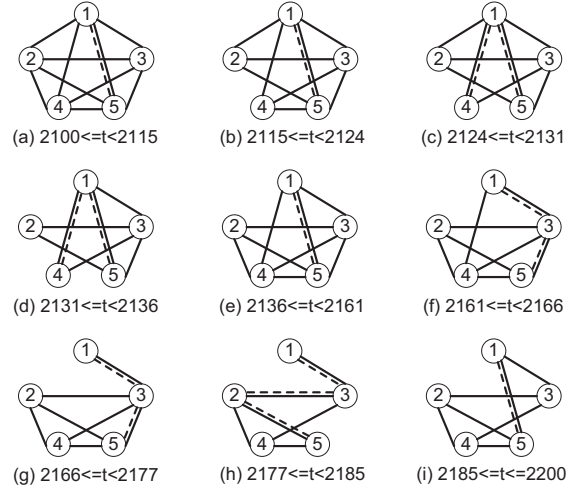


Figure 7. Typical RWM connectivity scenario for five node network.

Fig. 8 shows the average number in system x_1^5 and the end-to-end delay of the packets from node 1 destined for node 5 during the time interval [2100 2200]. Initially, every packet goes through the direct link (1-5). Then, during the time interval $2124 \leq t < 2136$ sec, a large transient increase in x_1^5 occurs due to traffic rerouting, when the link (4-5) breaks, the traffic x_4^5 needs to go through node 1 to reach the destination. This event also results in the increase of end-to-end delay of the packets x_1^5 , because of the higher utilization of node 1. At $t = 2136$ sec link (4-5) recovers and the x_4^5 traffic is rerouted back to the direct link (4-5). Starting from $t = 2161$ sec, link (1-5) breaks and the routing protocol redirects the packet x_1^5 through node 3 to node 5. Notice that at $t = 2177$ sec the link (3-5) is disconnected and it causes the x_1^5 traffic to take one more hop from node 3 to node 2 resulting in a further increase in the end-to-end delay. At $t = 2185$ sec the direct link (1-5) is working again and the traffic is rerouted over the direct link resulting in a decrease in the delay.

The fluid flow numerical results presented in Fig. 8, are calculated by assuming “homogeneous traffic” instead of (13) (14). For the lower bound, all sources are assumed to have homogeneous traffic with the period of $D_{max} = \max\{D_i\} = 1/0.16 = 6.25$, while for the upper bound, the homogeneous traffic becomes $D_{min} = \min\{D_i\} = 1/0.25 = 4$. Following the procedure in Section II.D case I, we obtain the utilization function for both bounds as $G_{lower}(x_T) = 0.0889x_T^3 - 0.4496x_T^2 + 1.0706x_T$, $G_{upper}(x_T) = 0.0743x_T^3 - 0.4159x_T^2 + 1.0993x_T$, which are used in (19) to form the two bounding models of the network. From Fig. 8, we notice that the lower and upper

bounds are tighter than ones in three node network scenario. Hence, the “homogeneous traffic” approach might be better choice than (13) and (14), when the $\sum_i^m N_i D_i / D/1$ queue has close arrival rates.

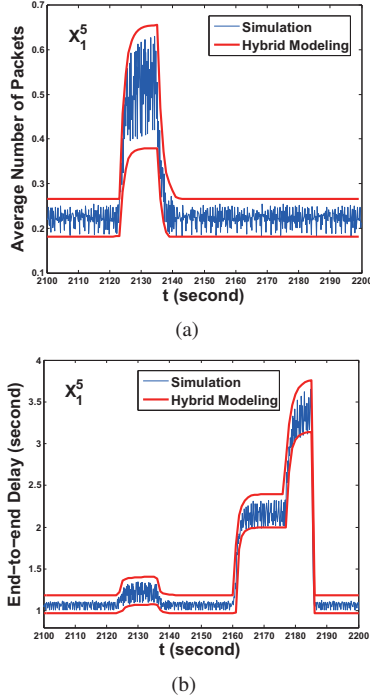


Figure 8. Average number of packets and end-to-end delay of traffic at node 1 destined for node 5.

4 COMPUTATION TIME COMPLEXITY

4.1 Complexity Analysis

The proposed hybrid modeling approach is composed of a fluid flow queueing model together with a time varying connectivity matrix. At each time interval, network metrics are evaluated by integrating a set of fluid flow model based differential equations with a specific connectivity matrix. Similarly to other areas of complexity theory, the exact number of arithmetic operations required for solving the differential equations with one step time by *Runge-Kutta* algorithm is hard to determine but an upper bound on the complexity order can be obtained.

Let dt refer to the time step size of solving the differential equations and T be the length of the time interval of interest. Then T/dt represents the number of total times to solve the set of differential equations. Let K denote the average time to execute one arithmetic operation on a CPU. Following [27], $C(n, p, \alpha)$ denotes the upper bound on the number of arithmetic operations required within one step

time, so that n differential equations can be solved by a p -th order explicit *Runge-Kutta* algorithm with maximum error $e^{-\alpha}$. According to the expression for $C(n, p, \alpha)$ in [27], with the pre-defined value of p and α , C increases linearly with n . As a result, only considering the varying of n in $K \cdot (T/dt) \cdot C(n, p, \alpha)$, the computation time complexity of our hybrid modeling TC_H is upper bounded by $O(n)$. For an M -node wireless network, regardless of the number of externally arriving traffic flows, the number of differential equations n is $M(M - 1)$. Therefore, the upper bound of TC_H is $O(M(M - 1))$.

4.2 Comparison by Sample Networks

To further evaluate the complexity we numerically determined the computational time for a set of sample network implementations. We constructed both OPNET discrete event simulation and the proposed hybrid fluid flow model for network settings of three nodes, four nodes, five nodes and thirteen nodes networks, as shown in Fig. 9. The numerical results are shown in Table I when both hybrid modeling and simulation were run on a laptop with Intel T7400 2.16GHz Duo-Core Processor with 2GB memory. For the discrete event simulation the time given is the total time to execute 4096 runs. This result shows that for the hybrid modeling, the numerical computation time is proportional to the number of differential equations. However, the computation time required to complete the OPNET discrete event simulation increases with the total number of events which is a complex function of the amount of traffic, topology changes and accuracy desired.

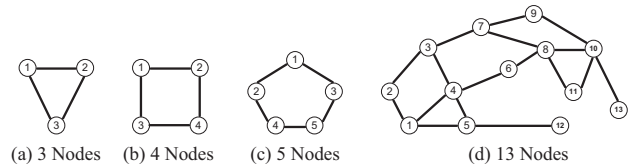


Figure 9. Topologies of sample networks

Table 1. Computation Time Comparison (in second)

# Nodes	# Flows of Externally Arriving	# Diff. Equation	Simulation (sec)	Hybrid Model (sec)
3	4	6	123.1	0.21
4	6	12	1173.7	0.48
5	8	20	11350.5	0.83
13	11	156	389723.4	6.62

5 Conclusion

In this paper, we propose a performance modeling technique to represent time varying behavior of multihop wireless networks with CBR traffic, using time varying connectivity matrix modeling and numerical method based queueing analysis. Network queues are modeled using fluid flow based differential equations and solved using numerical integration routines, while topology change is integrated into the connectivity matrix using deterministic or probabilistic modeling techniques. The proposed hybrid modeling approach can generally be applied to a wide range of queueing systems. Numerical results using the proposed model have been given in comparison with results from traditional discrete event simulations. The computation time required by both approaches is also shown side by side. We believe that this hybrid modeling approach is a proper tool for evaluating the time-varying behavior of multihop wireless networks. With the computation time saved from the fluid flow based hybrid modeling method, it is a tremendous gain in flexibility for modeling complex networks. Future work includes validating the model with testbed measurement results and developing additional features (e.g., multi-rate links) to increase the fidelity of the model.

References

- [1] I. F. Akyildiz and X. Wang, "A survey on wireless mesh networks", *IEEE Communications Magazine*, vol. 43, no. 9, pp. S23-S30, September 2005.
- [2] H. Moustafa and Y. Zhang, *Vehicular networks, Techniques, Standards, and Applications*, CRC Press, 2009.
- [3] I. Chlamtac, M. Conti, and J. Liu, "Mobile Ad Hoc Networking: Imperatives and Challenges", *Ad Hoc Networks(Elsevier)*, vol. 1, no. 1, pp. 13-64, 2003.
- [4] D. Logothetis and K. Trivedi, "The Effect of Detection and Restoration Times for Error Recovery in Communications Networks", *Journal of Network and Systems Management*, vol. 5, no. 2, pp. 173-196, 1997.
- [5] D. Tipper, J. Hammond, S. Sharma, A. Khetan, K. Balakrishnan and S. Menon, "An Analysis of the Congestion Effects of Link Failures in Wide Area Networks", *IEEE Journal on Selected Areas in Communications*, vol. 12, no.1, pp. 179-192, 1994.
- [6] M. Rumsewicz and D. Smith, "A Comparison of SS7 Congestion Control Options During Mass Call-in Situations", *IEEE/ACM Transactions on Networking*, vol. 3, no.1, pp. 1-9, February 1995.
- [7] D. Tipper, C. Charnsripinyo, H. Shin, and T. Dahlberg, "Survivability Analysis for Mobile Cellular Networks", in *Proc. Communication Networks and Distributed Systems Modeling and Simulation Conference* San Antonio, Texas, January 2002.
- [8] D. Tipper, Y. Qian and X. Hou, "Modeling the Time Varying Behavior of Mobile Ad-Hoc Networks", in *Proc. 7th ACM International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, October 2004, pp. 12 - 19.
- [9] W. Lovegrove, J. Hammond, and D. Tipper, "Simulation Methods for Studying Nonstationary Behavior of Computer Networks", *IEEE Journal on Selected Areas in Communications*, vol. 8, no. 9, pp. 1696-1708, 1990.
- [10] *Glomosim Tool*, <http://pcl.cs.ucla.edu/projects/glomosim/>.
- [11] *OPNET Simulation Tool*, <http://www.opnet.com/>.
- [12] *NS-2 Simulation Tool*, <http://www.isi.edu/nsnam/ns/>.
- [13] *QualNet Simulation Tool*, <http://www.qualnet.com/>.
- [14] *BonnMotion, Mobility scenario generation and analysis tool* <http://net.cs.uni-bonn.de/wg/cs/applications/bonnmotion/>.
- [15] T. Camp, J. Boleng, and V. Davies, "A Survey of Mobility Models for Ad Hoc Network Research", *Wireless Communication & Mobile Computing (WCMC): Special issue on Mobile Ad Hoc Networking: Research, Trends and Applications*, vol. 2, no. 5, pp. 483-502, 2002.
- [16] T. Lin and S. F. Midkiff, "Mobility versus Link Stability in the Simulation of Mobile Ad Hoc Networks", in *Proc. Communication Networks and Distributed Systems Modeling and Simulation Conference (CNDs)*, January 2003.
- [17] S. K. Hwang and D. S. Kim, "Markov Model of Link Connectivity in Mobile Ad Hoc Networks", *Telecommunication System*, vol. 34, no. 1-2, pp. 51-58, February 2007.
- [18] D. Tipper and M. K. Sundareshan, "Numerical Methods for Modeling Computer Networks under Nonstationary Conditions", *IEEE Journal on Selected Areas in Communications*, vol. 8, no. 9, pp. 1682-1695, December 1990.
- [19] W. Wang, D. Tipper and S. Banerjee, "A Simple Approximation for Modeling Nonstationary Queues", in *Proc. IEEE INFOCOM*, vol. 1, 1996, pp. 255 - 262.
- [20] M. J. Neely, and C. E. Rohrs, "Equivalent models and analysis for multi-stage tree networks of deterministic service time queues", in *Proc. the 38th Annual Allerton Conference on Communication, Control, and Computing*, October 2001.
- [21] M. J. Neely, "Exact Queueing Analysis of Discrete Time Tandems with Arbitrary Arrival Processes", in *Proc. IEEE International Conference on Communication (ICC)*, vol. 4, 2003, pp. 2212-2225.
- [22] M. J. Neely, "Equivalent Models for Queueing Analysis of Deterministic Service Time Tree Networks", *IEEE Transaction on Information Theory*, vol. 51, no. 10, pp. 3576-3584, 2005.
- [23] J. W. Roberts, and J. T. Virtamo, "The Superposition of Periodic Cell Arrival Streams in an ATM Multiplexer", *IEEE Transaction on Communications*, vol. 39, no. 2, pp. 298-303, 1991.
- [24] Elise T. Lee and John Wang, *Statistical Methods for Survival Data Analysis*, New York: John Wiley and Sons, 1992.
- [25] Piet Van Mieghem, *Performance Analysis of Communications Networks and Systems*, Cambridge University Press, 2006.
- [26] J. T. Virtamo and J. W. Roberts, "Evaluating Buffer Requirements in an ATM Multiplexer", in *Proc. IEEE GLOBECOM*, vol. 3, November 1989, pp. 1473 - 1477.
- [27] A.G. Werschulz, "Computational complexity of one-step methods for systems of differential equations", *Math. Comput.*, vol. 34, no. 149, pp. 155-174, 1980.