Abstract— Mobile ad hoc networks suffer from route breakdowns caused by nodal mobility. In a network that aims to support critical interactive real time data and flows, as well as the un-interrupted transport of a critical series of messages, it is essential to identify robust routes that can be used for the un-interrupted transport execution of such transactions. Noting that route failures can induce long re-routing delays that may be highly inter-ruptive for many applications and message / stream trans-actions, it is beneficial to configure the routing scheme to send a flow across a route whose lifetime is longer, with sufficiently high probability, than the estimated duration of the activity burst that it is selected to carry. We evalu-ate the ability of a mobile ad hoc wireless network to distri-bute flows across robust routes by introducing the robust-throughput measure as a performance metric. For example, only transactions that are completed without being prematu-remently interrupted may convey data to their intended users that is of acceptable utility and is thus accounted for. We describe the mathematical calculation of a network’s robust through-put measure, as well as its robust throughput ca-pacity. In order to transport flows in mobile ad hoc wireless network in a robust fashion, we introduce the Robust Flow Admission and Routing algorithm (RFAR). Under this on-demand routing scheme, during the route discovery phase, nodal routers configure routing forwarding entries to for-ward flow packets only across links that induce a sufficiently high cumulative route robustness level. Shortest (or least end-to-end delay) such routes are subsequently discovered and established, serving to best utilize network capacity or resources while meeting flow robustness objectives. We de-monstrate through mathematical analysis and by using simulation evaluations that such a robust routing mecha-nism serves to significantly enhance the robust throughput performance of mobile ad hoc wireless network systems.

I. INTRODUCTION

The performance of a mobile ad hoc wireless network is impacted by the dynamic stochastic process characteristics of its underlying links (and the associated noise interfer-ences, data rates, ranges, communications capacity levels), nodes (e.g., their mobility patterns and resource states), the underlying graph connectivity of the network topology, and the application induced traffic loading processes and their required quality of service (QoS) objectives. Under typical on-demand ad hoc routing algorithms, a source node that wishes to communicate across the network, initi-ates a route discovery process. Consequently, a route may be discovered for the transport of messages generated as part of the underlying session flow. Such routes are perma-nently established and periodically updated under a proac-tive routing scheme. In either case, the robustness of the route is generally not involved as a requirement for its selec-tion. Consequently, route break-ups will frequently occur, induced by nodal mobility and/or nodal and link failures as well as by fluctuations in the communications transport quality experienced across the network’s communications links, induced by signal interferences, fading and multi-path phenomena and other causes producing ambient and environmental noise and signal interference processes. As a result, a flow’s route used to transport a file, or a group of packets that are part of an end-to-end user transaction, may be broken even before the flow’s transaction (or cor-responding session or call holding time) has expired.

When an active route is broken, the network on-demand routing mechanism will act to find alternate routes to re-route the impacted flows. Under a proactive routing scheme, link or node failures induce routers to calculate new forwarding entries that are then used to re-route af-fected flows. In either case, due to the typical relatively slow reaction times characterizing the control mechanisms involved in establishing new routes for mobile ad hoc wire-less network systems, relatively long re-routing delays can be experienced. Such latencies may be readily perceivable at the impacted layer (including often also the involved ap-plication function), and can cause significant degradation in the quality of the provided transport process for many applications. Thus, induced by such perceivable re-routing delays, the reception of data packets embedded in flows that use routes that are prematurely terminated may be of reduced value to their intended receivers. An applica-tion may generate a program (file, collection of messages) that must be delivered, as a whole, across the network in a critically timely manner. The re-routing process tends often to be unacceptably interruptive for the networks under-consideration, it is essential to ensure that the selected route is sufficiently robust. We thus often need to ensure that the underlying transactions are executed in full, with high probability, without any (including re-routing) interrup-tion.

The commonly used performance index of network throughput represents the average amount of data that is delivered to intended destinations within a prescribed pe-roiod of time. Often, a ‘goodput’ measure is used to evaluate the throughput rate of a networked system, or across spe-cific layers of such a system, such as the transport layer. Packets that cannot be correctly decoded at their intended destinations, or that are received out of order or dupli-cated may not be included in the total packet count used to produce the underlying ‘good’ throughput rate. For ex-ample, in [1, 2], the ‘goodput’ performance of mobile ad hoc networks is evaluated, assuming this ‘good’ throughput measure not to account for packets that are received incor-rectly or out of order. Clearly, either of these measures, do not characterize the ability of a network to successfully transport complete (multi packet) files, or complete un-interruptedly the transport of whole transactions, as noted above, a flow’s transaction may be prematurely interrupted due to the breakup of its allocated route. To the best of our knowledge, the ability of a mobile ad hoc wireless net-work to transport whole files / programs, or execute full ses-sions, in an uninterrupted fashion, ensuring that with high probability the flow’s route will stay intact for the du-
ration of the transaction, has not yet been characterized and studied. This is the objective of this paper.

Clearly, executing a transaction (without a preceding control phase, or with it, such as the one that is occurring as part of an admitted flow or an established session) across the network may involve the transport of a large number of packets. In this case, the correct reception of a fraction of these packets within a prescribed time delay window, or the delay experienced (impacted by relatively long re-routing latencies) by a receiving entity that waits for the complete (and ordered) set of packets to be received, may reduce significantly the value of the transport service provided by the network. Such interruptions are caused by noise processes and interferences experienced across each communications link as well as by mobility induced breakups of links and routes. While the earlier have been considered before (inducing, for example, the calculation of 'goodput' measures that exclude incorrectly received packets, as noted above), we account here also for the latter. Clearly, such an issue is of prime importance for mobile ad hoc wireless network systems for which link breakdowns tend to occur routinely.

We evaluate the ability of a mobile ad hoc wireless network to distribute flows across robust routes by introducing the robust throughput measure as a performance metric. The latter describes the robust data rate level received by the intended destinations. For example, only transactions that are completed without being prematurely interrupted may convey data to their intended users that are of acceptable utility and are thus accounted for. We describe the mathematical calculation of a network’s robust throughput measure, as well as its robust throughput capacity.

To improve the robust throughput features exhibited by a networked system, we present the robust flow admission and routing algorithm (identified as RFAR). As an on-demand ad hoc routing scheme, the flow admission control scheme is used by nodes to reject, during the route discovery phase, a route request message for which it is not possible to discover a route that meets the robustness requirement desired for the indicated application. This algorithm is shown to yield an operation that maximizes the robust throughput behavior of the system. We present mathematical models for the analysis and design of such robust network systems, as well as carry out extensive simulations that confirm the effectiveness of our on demand routing and flow control schemes in ensuring the system with high robust throughput performance.

The rest of the paper is organized as follows. The robust throughput measure and the system’s attainable throughput and robust throughput capacity levels are defined and discussed in section II. In Section III, we present the robust flow admission and routing algorithm. We present mathematical formulas that can be employed by the system designer to evaluate the robust throughput behavior of the underlying network system. In Section IV, we use our mathematical formulas, as well as network simulation evaluations to exhibit the robust throughput performance behavior of network systems. We confirm the precision of our analytical model and demonstrate the robust throughput performance upgrades that can be attained through the use of our robust flow admission and routing protocol. Conclusions are drawn in section V.

II. ROBUST THROUGHPUT

Let \( X = \{X_t; t \geq 0\} \) be the system’s session size process. It is a stochastic process, whose state space is \( E = \{0, 1, 2, \ldots\} \), that represents the temporal evolution of the system’s total number of supported flows (or sessions). (See Fig. 1 for an illustrative realization of such a process and of the associated reward variables defined below.)

The random variable \( X_t \) denotes the number of sessions supported by the system at time \( t \). The \( n \)-th supported session is admitted into the network system at time \( A_n \) and is assumed to be terminated (and thus departing from the system, either by interruption or due to its successful completion) at time \( R_n \). We proceed to define measures of system performance that depend on the nature of the service provided to the underlying supported sessions. For this purpose, we assume that the proper nature of the service provided to a session (or flow) can be fully determined once the session has departed. For example, if the latter departure event represent a successful completion of the transport task associated with this session (e.g., the session aims to transfer a program file to the destination and is regarded to successfully accomplish its task only if upon its departure the full program has been correctly transferred), we proceed to provide the system with full credit for successfully accomplishing the underlying transport mission. In turn, consider situations under which partially delivered data files are held only for a short period of time and are then discarded. In this manner, any later resumption of the transport of messages that belong to the underlying original program/document will have to involve the retransmission of messages sent during the lifetime of a connection that has prematurely failed. In this case, if the departure of a session is triggered by an interruption event (induced, for example, by breakdowns occurring along the session’s route due to nodal mobility or due to link transmission fading events or nodal failures), the task to be achieved by the underlying transfer associated with this session is regarded to have failed. In this case, no credits should be awarded to the system for the (partially completed) transport actions executed by the network. Of course, partial credit (payoffs or rewards) are often appropriate, depending upon the definition of the conditions required in declaring for successful transfer / transport. Such determinations can be made at session departure time.

Accordingly, for departing session \( n \), let random variable \( Y_n \) denote the payoff that the system obtains at time \( R_n \) upon the session’s departure. The corresponding reward process is denoted as \( Y = \{Y_n; n = 1, 2, \ldots\} \), with the rewards assuming values in \( \mathbb{R} = [0, +\infty) \).

The total reward gained by the system over the pe-
period of operation $[0, t]$ is represented by random variable $Z_t = \sum_{n=1}^{N^p} Y_n$, where $N^p$ denotes the number of session departures occurring over the period $[0, t]$. The average reward gained by the system per unit time, denoted as $\eta$, is an indicator of the operational efficiency of the system:

$$\eta = \lim_{t \to \infty} \frac{E(Z_t)}{t}.$$

For example, when the gained reward process $Y = \{Y_n\}$ is represented as a sequence of independent identically distributed (reward) random variables, each assuming a fixed mean value, and is also statistically independent of the session departure point process (and associated counting process; or more generally assuming that the variable $N^p$ is a stopping time relative to sequence $Y$ so that Wald’s Lemma can be applied, assuming also a finite value for the average number of departures that take place over $[0, t]$) we write

$$\eta = \lim_{t \to \infty} \frac{E(\sum_{n=1}^{N^p} Y_n)}{t} = \lim_{t \to \infty} \frac{E(N^p_t)E(Y)}{t}.$$  

The session departure rate is defined as: $\lambda_D = \lim_{t \to \infty} \frac{E(N^p_t)}{t}$; hence,

$$\eta = \lambda_D E(Y).$$

Assume the session arrival (offered) rate and session (flow) blocking ratio (or probability) are denoted as $\lambda_c$ and $P_B$, respectively, then $\lambda_D = \lambda_c(1 - P_B)$. Thus,

$$\eta = \lambda_c(1 - P_B) E(Y).$$

A. The network’s throughput rate

The network system’s throughput is commonly defined as a measure that expresses the (limiting) average number of information units delivered by the network to their corresponding intended destination nodes, per unit time. To calculate the system’s throughput rate by using the expressions presented above, we set the $n$-th session reward variable $Y_n$ to be equal to the total amount of information units transported by the $n$-th departing session across the network, independently of the nature of the termination event associated with the departure of this session. Hence, we set $Y_n$ represent the cumulative total number of information units of session-$n$ ‘successfully’ delivered to the destination node(s) during the realized lifetime of the session (independently of the nature of the termination event associated with the departure of this sessions). We express it as $Y_n = F_n H_n$, where $F_n$ and $H_n$ denote the information data rate (realized for the end-to-end transfer of information for session $n$) and the realized duration of session $n$, respectively. The value assumed by $H_n$ is affected by two possible events. When the session is successfully completed, $H_n = T_n$, where $T_n$ denotes the intended session’s holding time. Otherwise, the session is pre-maturely terminated, so that $H_n = B_n$, where $B_n$ represents the length of time expired till termination (due to route breakup) for session $n$. Therefore, we set $H_n = \min\{T_n, B_n\}$.

The system’s throughput rate ($f$) is thus expressed as

$$f = \eta = \lambda_c(1 - P_B) E(F H) = \lambda_c(1 - P_B) F E(H). \quad (1)$$

Assuming (for simplifying the expression, with no loss in generality) the data rate level (measured in bps) process

$$F = \{F_n; n = 1, 2, \ldots \}$$

and the session holding time process

$$H = \{H_n; n = 1, 2, \ldots \}$$

to be mutually statistically independent; and the corresponding average levels are denoted as $E(F)$ and $E(H)$, respectively.

In [3], we have shown that the session blocking probability can be calculated by modeling the network as an $M/M/m/m$ queueing system. For this system, link flows are modeled as customers. The link flow arrival rate is equal to $E(\Pi)\lambda_0$, noting that each flow contributed an average of link flows, where is the path length process of the flow sessions. The service rate of each server is given by $1/E(H)$. The number of available service channels $m$ is set equal to the level that will yield an overall aggregate service rate that is equal to the aggregate transport capacity of the network. Hence, we have:

$$m = \left[\frac{SRF \cdot R}{E(F)}\right],$$

where SRF denotes the average value of the spatial reuse factor [4]. This factor describes the average number of simultaneous link layer transmissions that are successfully received by their respective (cross link) receiving nodes (conducted over the area of network operations, under the employed cross-layer transmission and networking schemes). The parameter $R$ denotes the data rate at which each transmitter operates. We note that for a wireless network in which each node employs an omni-directional antenna, the underlying layer communications channel is a multiple access channel that is shared by neighboring nodes. In this case, nodes that are located too close to each other (or, rather to the intended receivers) may not be able to conduct simultaneous transmissions since the latter may cause too high interference signal levels at the corresponding receivers. In [5], we have shown that the SRF can be estimated by disk-covering approach thus $SRF = \frac{A}{\pi r^2}$, where $A$ is the total area of the network and $r$ is the communication range between two mobile nodes.

Using Erlang’s Loss formula for this queueing system, we obtain the link flow blocking probability to be given by

$$P_B = P(X = m) = \frac{(E(\Pi)E(H)\lambda_0)^m}{m!\sum_{i=0}^{\infty}(E(\Pi)E(H)\lambda_0)^i/i!}.$$

B. The network’s robust throughput rate

As illustrated above, for certain applications, it is essential that the flow’s session be carried out to completion to derive full benefit from the underlying interaction or transport of information. In such cases, a disruption, or premature termination, of the communications process in the midst of a session can significantly discount the value of packets delivered to the destination prior to such an early termination event.

Consequently, it is often the case that the effective value of the delivery of information units (such as messages/packets) to the destination is not realized unless, and until, the transport session is carried out without interruption to completion. For example, the delivery of an executable computer program file across a network to a destination node is of no value to the destination if the transport process for this flow is prematurely terminated, though a partial delivery of certain packets may be successfully completed. In such a case, the network throughput realized by the delivery of packets that represent just a portion of the
transaction should not be assigned any credit. For such applications, a transaction is either given a full throughput (constructive performance) credit (if it is executed to completion) or no credit at all. Yet for some applications, it is appropriate to represent the benefit gained by the end-users of the network by including, in calculating a measure of end user performance, partial credit for transactions that result in the network system delivering to the end users incomplete portions of the loaded message units.

We thus define a measure that accounts for benefits gained by the end-users involved from the reception of completed or partially completed transactions by introducing the Robust Throughput (RT) metric.

In this case, we let $Y_n = F_n H_n S_n$, where $S_n \in [0, 1]$ represents the fraction of the credit that session $n$ obtains upon completion, and $S = \{S_n; n = 1, 2, \ldots\}$. We define the efficiency factor associated with such a defined reward function as the system’s Robust Throughput (denoted as $f_r$). That is,

$$f_r = \eta = \lambda_0 (1 - P_B) E(F) E(HS).$$

When $S_n = 1$ for all $n$, every session receives full credit for all the data that has been transported during the lifetimes of its route, independently of the cause of session termination and of the realized duration of session time to termination. In this case, the Robust Throughput $f_r$ metric measures the system’s throughput level.

Consider next the special case under which every session receives full credit for the transport of its packets only when the intended transport is successfully completed (or equivalently, when the session is said to be successfully completed), and no credit is awarded when the session is prematurely terminated. The binary random variable $S_n$ is set equal to 1 and 0, respectively. We often identify such a robust throughput measure $f_{r-CB}$ as the network system’s Completion-Based Robust Throughput metric. We have:

$$f_{r-CB} = \lambda_0 (1 - P_B) E(F) E(HS)$$

$$= \lambda_0 (1 - P_B) E(F) E(H|S = 1) P(S = 1) + E(H|S = 0) P(S = 0)$$

$$= \lambda_0 (1 - P_B) E(F) E(H|S = 1) P(S = 1)$$

$$= \lambda_0 (1 - P_B) E(F) E(T|S = 1) P(S = 1).$$

When flow is successfully terminated ($S = 1$), we have, $H = T$. Thus, $E(H|S = 1) = E(T|S = 1)$. The Completion-based Robust Throughput metric is a measure that represents the capability of a network system to transport message flows without incurring a high rate of premature transaction terminations. We note that a network that yields a high throughput level doesn’t necessarily provide a high robust throughput value.

C. Throughput and robust throughput capacity

The network system’s throughput capacity $f^*$ is defined as the maximum throughput level that a network can sustain, under prescribed operational and traffic loading conditions.

For example, the network admission flow control mechanism used can be designed to ensure the desired level of per flow service provided by the network. Once the network loading, admission control scheme and other networking procedures are applied, an overall throughput rate $f[\text{bps}]$ is attained. The total internal traffic rate (being the aggregate sum of all the link traffic rates experienced within the network under the latter throughput rate) observed under the latter conditions is denoted as $f_l[\text{bps}]$. Clearly, the following relation holds:

$$f^* = \frac{f_l}{E(U) E(H)}. \tag{3}$$

From Eq.(1), when the offered rate $\lambda_0$ is very high, the network system attains its maximum throughput. Therefore,

$$\lambda_0 (1 - P_B) E(F) E(H) E(I) = R \cdot SRF.$$

Similarly, use Eq.(2), we have the maximum robust throughput given by,

$$f_{r-CB} = \frac{R \cdot SRF \cdot E(HS)}{E(U) E(H)},$$

under prescribed loading profiles and operational conditions (including transmission, networking schemes). Specifically, for completion-based robust throughput, the maximum robust throughput is given by,

$$f_{r-CB} = \frac{R \cdot SRF \cdot P(S = 1) E(T|S = 1)}{E(U) E(H)}.$$

As noted above, to calculate the network’s completion based robust throughput metric, we must calculate the probability $P(S = 1)$ that expresses the rate at which a flow’s (session’s) route remains operational for a period of time that is not shorter than the session’s duration (or, equivalently, the time that it takes to complete the involved flow transactions, or corresponding transport of packets). For this calculation, we need to statistically characterize the lifetime of an intact (uninterrupted) network route, for routes of different lengths (number of hops). We carry out such analysis in the following section.

D. Survival Time of Links and Routes

In [6], we have examined the behavior of link and route lifetimes by focusing on breakups that are induced by nodal mobility. Assuming a random waypoint mobility model (with relatively low values assumed for the times spent by nodes in pausing at the area boundary), we have shown that the distribution of the route survival time is well approximated by an exponential distribution. It is thus written as,

$$P(T_m \leq t) = 1 - e^{\kappa t},$$

where $\kappa$ is a constant that is determined by the mobility pattern of the nodes. We have shown the parameter of the underlying link lifetime distribution to be well approximated by setting $\kappa = \mu \nu_1 + \nu_2 r$, where $\mu$ is a parameter determined by the mobility pattern (see [6]), $r$ denotes the link’s communications range, and $\nu_1, \nu_2$ represent the speeds of the underlying link’s end nodes; so that we have, for $t \geq 0$:

$$P(T_m \leq t) = 1 - e^{-\frac{\mu \nu_1 + \nu_2 r}{t}}.$$

To represent link failure events, we assume the following model. A link breakup can be induced by either one of the following two factors: (1) Nodal mobility that sets the link’s end nodes to be at a distance that exceeds the threshold level, and thus making communication ineffective at
the desired bit error rate level. (2) Link outages that occur when the nodes are located within the designated communications range ($r$). Such outages can be caused by noise and interference processes, as well as the mobile character of the end nodes.

For mathematical simplicity, we assume the link’s time to fade ($T_f$) that represents the above mentioned second factor to also follow an exponential distribution, so that for $t > 0$ we have:

$$ P(T_f > t) = e^{-\mu_f t} = e^{-\frac{t}{T_f}}, $$

where $T_f = 1/\mu_f > 0$, denotes the average time to such an outage occurrence.

For illustrative purposes, assume the above mentioned two lifetime periods to be statistically independent (for scenarios under which link breakup events caused by nodal mobility are approximately independent of outage causing fading phenomena). In this case, when we combine these components, the link life time ($T_l$) is characterized by the following exponential distribution, for $t > 0$:

$$ P(T_l > t) = P(T_m > t)P(T_f > t) = e^{-\beta t}e^{-\mu_f t} = e^{-\left(\frac{\mu_f}{\beta} + \frac{1}{T_f}\right)t}. $$

We have also showed in [6] that a good approximation for the survival time of a flow’s route, for the random mobility model under consideration, is obtained by assuming the variables representing the lifetimes of the links that make the route to be statistically independent. Consequently, noting that a route will break as soon as one of its links fails, the distribution of a route’s survival time is calculated as, $t \geq 0$,

$$ P(T_m < t) = 1 - P(T_m > t) = 1 - \min \{P(T_m^{-1,i} > t)\} = 1 - \prod \{P(T_m^{-1,i} > t)\} = 1 - e^{-\mu_t(v_0 + 2v_1 + \ldots + 2v_{m-1} + v_e)t/r}. $$

For aid in the operation of the routing scheme to be introduced in a later section, we denote the weighted sum of the nodal speed along the route as, $w = v_0 + 2v_1 + \ldots + 2v_{m-1} + v_e$. When we include link fading and shadowing effects we can write

$$ P(T > t) = e^{-\left(\frac{w}{r} + \frac{1}{T_f}\right)t}. $$

### III. Robust Flow Admission and Routing Algorithms

In this section, we study routing and flow admission schemes that minimize the transport capacity level required for routing flows to their intended destinations. Consequently, such schemes yield desirable delay-throughput behavior. In order to guarantee the robustness of selected routes, the scheme must admit only flows that can be ensured to experience: (1) The desired robustness target level and (2) desired end-to-end packet delay (mean and jitter) level.

To achieve such an operation, we use the following distributed on-demand routing and flow admission (RFAR) scheme: (1) During the route discovery period, as integral part of the route discovery process, involved nodes choose to forward a RREQ message towards its destination node only if it determines that hence-to-for route traveled by the RREQ message (including the outgoing link) is sufficiently robust, in relation to the end-to-end robustness level desired for the underlying flow. (2) To maximize the system’s robust transport capacity level (RTC) level, the destination node of a flow’s flooded RREQ messages proceeds to select the shortest path. For this purpose, the latter node waits a period of time to receipt of flow RREqs and then selects among them in accordance with an employed index. The latter can be based on the use of weighted delay and path length and/or robustness level metrics.

The route robustness requirement index (RRRI) of a flow, denoted as $\beta$ for a flow of application class c (e.g., telemetry, voice, interactive critical data, file transfer, etc.) is defined as the maximum value of the probability that the flow’s route will be disconnected before the flow’s session (transport) is completed.

To meet a flow’s route’s robustness objective level, each node performs the following comparison in determining whether to forward a RREQ message for establishing a path for a flow of class c (noting that the RREQ carries in its header data regarding the cumulative robustness index)

$$ w + \frac{r}{\mu} \leq \frac{r \ln \beta}{\mu}. $$

That is, in order to guarantee a prescribed robustness index level, the weighted sum of the speeds of the nodes along the route $w$ must be limited by route robust requirement threshold (RRRT) $\alpha = \frac{r \ln \beta}{\mu}$. That

We describe our mobile ad hoc robust flow admission and routing algorithm (RFAR) as follows. When a source node desires a route to a destination for which it does not already have a route, it broadcasts a route request (RREQ) packet across the network. The RREQ contains the following: (1) source node’s ID; (2) requested flow parameter, (3) current sequence number, (4) broadcast ID; (5) route robust requirement value (RRRV). RRRV characterizes the robustness of the prescribed testing route is up to this node. Nodes receiving this packet do the following steps:

1. The node updates its information for the source node and set up backwards pointers to the source node in the route tables.

2. The node checks if such request has already been processed by itself. If so, it discards the RREQ and does not forward it.

3. The node checks the (a) congestion constraints and (b) robustness constraints. If it doesn’t meet the constraints the node discards the RREQ and does not forward it. The current nodal congestion state (as determined, for example, by the nodal queue length monitored over a recent sliding window period) is determined and compared with a prescribed threshold level. The robustness level of the outgoing link is estimated over a recent sliding window period, by monitoring the link robustness index (defined as the probability that a link will not fail or break over a period that is equal to a single time unit (e.g., 1 sec) and possibly incorporate various fading and noise conditions); this value is used to aggregate the robustness value included in the packet’s robustness field. The aggregated robustness level is then compared with the robustness threshold. The
node modifies the received RRRV by
\[
RRRV_{\text{new}} = RRRV_{\text{old}} + \frac{1}{2}(v_i + v_j) + \frac{r}{\mu L}\]
where \(v_i\) and \(v_j\) are the speed of the node itself and the node who it receives RREQ from. The node then compares the new RRRV with RRRT. If the new RRRV is greater than RRRT, the RREQ is discarded. Otherwise, it then may send a route reply (RREP) back to the source if it is either the destination or if it has a route to the destination with corresponding sequence number greater than or equal to that contained in the RREQ. Otherwise, it rebroadcasts the RREQ to its neighbors. As the RREP propagates back to the source, nodes set up forward pointers to the destination. Once the source node receives the RREP, it may begin to forward data packets to the destination.

In effect, the RFAR scheme identifies the shortest path that satisfies the robustness and packet delay performance requirements.

In comparison, we denote the routing and flow admission algorithm that do not impose robustness target measures (such as AODV [7], DSR [8], MBN [5,9]) as Non-RFAR (NRFAR). In NRFAR, the scheme doesn’t check the robustness constraints of given routes. Effectively, it omits step 3(b) above.

IV. SIMULATION AND PERFORMANCE EVALUATION

We have used the models presented above as well as run simulations of various networking scenarios for the purpose of studying the robust throughput behavior of network systems. In particular, we aim to study the relationship between a system’s throughput and robust throughput performance levels realized by network systems, as they relate to the underlying key system parameters such as nodal speeds, communications link range, and network traffic loading scenarios.

For this purpose, we have run a simulation of an ad hoc wireless network over an area of operation of size 1500m × 1500m, a link maximum communications range level of \(r = 250m\), and a channel data rate of 1 Mbps. To focus on the performance of the system at the network layer, we assume a MAC layer operation that schedules packet transmissions across the wireless channel in such a manner that active nodes within 2 hops from each other are not scheduled to time simultaneously transmit their packets. We vary the speed of the mobiles. Nodes have been assumed to move in accordance with a Random Waypoint mobility model [10] with a pause time equal to zero. Flows are initiated at their source nodes in accordance with a Poisson arrival process at the prescribed network-wide flow arrival rate. The flow’s session duration is prescribed to follow an exponential distribution with mean holding time equal to \(T\). For each flow, we assume that the source-destination pair is randomly chosen among all 500 network nodes. We use the results of the simulations to confirm our analytical models and to compare the performance behavior of the system, under the use of the Non-Robust Flow Admission and Routing algorithm (NRFAR) and the Robust Flow Admission and Routing algorithm (RFAR), as expressed by the attained throughput, robust throughput and route breakup rates.

For performance comparison purposes, we have examined the following three operational scenarios: (1) All nodes move at fixed speed of 5m/s. (2) Half of the nodes move at fixed speed of 5m/s, while the other half move at a fixed speed of 10m/s. (3) Nodes move at fixed speed, the speed of which is independently for each node in accordance with a uniform distribution, ranging from 0 to 10m/s. For each scenario, we compare the performance realized by employing the NRFAR and RFAR (under a prescribed robustness factor tail of 95%) algorithms.

We demonstrate that a network system operation that yields high throughput levels does not necessarily provide a high robust throughput behavior. We show that using our robust admission control scheme, the network is capable of achieving a highly effective operation that yields a high level of robust throughput.

Clearly, when network communications capacity resources are highly limited, it is better to admit only flows that can be accommodated across routes that offer the desired levels of robustness and packet delay behavior. Our dynamic congestion and robustness control schemes, implemented as part of the route discovery control process, act to ensure that only such flows are admitted.

We note that for the illustrative cases used here, the network has been loaded by flows that span the overall area of operation so that the SRF estimate that was calculated based on disk covering over the whole area is noted to yield excellent fit.

In Figure 2, we show the system’s throughput vs. offered rate performance curves. We note that, as the offered load is increased, the throughput level attained under the use of the NRFAR and RFAR algorithms asymptotically approaches the network’s throughput capacity level. The value of the latter is noted to be also determined by using Eq.(3). The throughput level achieved under the use of the NRFAR scheme is lower by about 20% that that attained by the RFAR procedure when to offered rate level is moderate (ranging from 1 to 4 Mbps). This is explained by noting that under the RFAR scheme, flows that cannot meet sufficient robustness constraints are blocked.

In Figure 3, we exhibit the robust throughput behavior of the network system under the use of the RFAR and NRFAR algorithms. We observe that as the offered loading rate is increased to a high level, the robust throughput level attained by the RFAR algorithm approaches to the system’s robust throughput capacity level, while that achieved by the NRFAR scheme is equal to only about
Figure 3. Robust throughput level for RFAR and NRFAR algorithms

<table>
<thead>
<tr>
<th>v</th>
<th>NRFAR</th>
<th>RFAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>27.323%</td>
<td>5.9735%</td>
</tr>
<tr>
<td>10</td>
<td>31.897%</td>
<td>5.8229%</td>
</tr>
<tr>
<td>5</td>
<td>46.372%</td>
<td>6.6233%</td>
</tr>
</tbody>
</table>

TABLE I
Route Breakup Rate Under the RFAR and NRFAR Algorithms with Desired Route Breakup Rate: < 5%

40% of the robust throughput capacity level. The robust throughput performance behavior is noted to degrade as the nodal speed levels are increased (namely at speed levels that are increased to 5 and 10m/s). This degradation is more pronounced for the operation conducted under the NRFAR scheme. Note for the all nodes speed equal to 5m/s case, RFAR decides to admit a flow based on the path length of the expected route.

In Table I, we show the route breakup rate incurred under each algorithm. The prescribed (desired) route breakup rate (used for admitting flows under the RFAR scheme) has been set to 5%. We observe that under the NRFAR scheme, route breakup rate levels ranging from 5.8% to 6.6% are realized. Under the NRFAR scheme, the route breakup rate is noted to be significantly higher, ranging from 27% to 46%. Clearly, at such a high route breakup rate, significant network capacity levels are used for route re-discovery attempts.

V. CONCLUSIONS

Robust throughput and robust throughput capacity indices are proposed in this paper as measures that serve to characterize the ability of a mobile ad hoc wireless network to provide highly survivable transport of flows for applications that require flow transactions to be carried out to completion to yield maximum benefit. We note that a network system that is designed to yield a high throughput rate doesn’t necessarily provide its users with a high measure of stability and consequently may be characterized by low robust throughput performance. To provide a robust throughput transport service for flows that involve multi-packet transactions that should not be prematurely disrupted (subjecting them to termination or to long re-routing delays), we present a new scheme that implements our robust flow admission and routing algorithm. To assess the robustness of a route, we employ a model that describes mathematically the lifetimes of links and routes in terms of underlying network system parameters that include nodal mobility values, communications ranges and transaction duration lengths. Our on-demand robust routing algorithm selectively discovers routes that can be probabilistically assured to survive for the duration of the underlying sessions or file traversals. As an alternative to use our formulas to estimate link robustness levels, we note that an implemented mobile entity can employ cross layer monitoring to dynamically (on a sliding window basis) estimate the stability and communications quality of its attached links. Interactions with its attached (cross link) nodes can also be effectively used to identify their SINR states and thus deduce a link quality and robustness measures. Also, to improve the robust behavior of a network, it is clearly advantageous to position stable and capable relay nodes, including UGVs and/or UAVs, in locations that serve to reduce the mobility impact of nodes on the stability of selected routes. The methods developed in this paper can then be used by the network designer to evaluate the robust throughput enhancement attained through such operations.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (NSF) under Grant No. ANI-0087148, and by University of California/Conexant MICRO Grant No. 04-100 and by University of California/ Nokia MICRO Grant No. 05-054.

REFERENCES


7 of 7