End-to-end vs. Hop-by-hop Transport under Intermittent Connectivity (Invited Paper)

Simon Heimlicher
heimlicher@tik.ee.ethz.ch

Merkouris Karaliopoulos
karaliopoulos@tik.ee.ethz.ch

Hanoch Levy *
hanoch@post.tau.ac.il

Martin May
may@tik.ee.ethz.ch

Computer Engineering and Networks Laboratory
ETH Zurich, Switzerland

ABSTRACT

This paper revisits the fundamental trade-off between end-to-end and hop-by-hop transport control. The end-to-end principle has been one of the building blocks of the Internet; but in real-world wireless scenarios, end-to-end connectivity is often intermittent, limiting the performance of end-to-end transport protocols. We use a stochastic model that captures both the availability ratio of links and the duration of link disruptions to represent intermittent connectivity. We compare the performance of end-to-end and hop-by-hop transport over an intermittently-connected path. End-to-end, perhaps surprisingly, may perform better than hop-by-hop transport under long disruption periods. We propose the spaced hop-by-hop policy which is found to dominate (in terms of delivery ratio) the end-to-end policy over the whole parameter range and the basic hop-by-hop policy over most of the relevant range.

1. INTRODUCTION

Emerging wireless communications paradigms face challenges that differ substantially from those in the wired Internet. Wired networks are in general well-connected over reliable links and network partitioning is not an issue; thus it is assumed that source and destination are continuously connected by an end-to-end path. Early simulation experiments with routing protocols for ad hoc networks [12, 18, 19] suggested that end-to-end connectivity can also be assumed in dense wireless multi-hop networks. The de-facto standard transport control in wired networks, TCP [4], was found to under-utilize the network in mobile wireless networks, and countless proposals of wireless-enhanced TCP modifications followed [3]. However, recent real-world measurement studies of wireless link capacity [1, 15] reveal frequent link disruptions even among stationary nodes, thus invalidating many assumptions of simulation studies. Moreover, node density is often sparse in real-world mobile wireless networks [13] and providing end-to-end connectivity is not feasible [14]. Even if an end-to-end path exists for some time, signal propagation impairments, interference, and node mobility incur frequent network partitioning. Thus, end-to-end connectivity in wireless multi-hop networks such as sensor, ad hoc, and mesh networks is often intermittent; we refer to such scenarios as *intermittently-connected networks (ICNs)*. In intermittently-connected networks, the performance of end-to-end transport protocols is limited since such protocols transfer data only while a connected path is available, leading to low network utilization.

Instead of proposing yet another modification to TCP, we wish to take a step back and *re-visit* the fundamental design decision of the transport layer. The original objective of the TCP/IP protocol suite [20, 21] was to interconnect heterogeneous networks. This objective required a transport layer implementation that made minimal assumptions about intermediate network elements and resulted in the end-to-end transport layer of the current Internet architecture.[1]

The alternative to an end-to-end transport layer is a distributed implementation along the network nodes that lie on the data transport path; we refer to such an implementation as *hop-by-hop transport*. The performance of hop-by-hop transport is superior to the end-to-end alternative in many aspects. However, this advantage comes at the price of higher protocol complexity and additional memory and processing requirements.

In intermittent-connectivity scenarios, hop-by-hop transport may transfer data along partial paths and deliver the same amount of data with fewer packet transmissions and lower latency. These advantages are particularly tempting in mobile networking applications, where a reduced number of link transmissions results in lower energy consumption and interference and also has a favorable impact on the capacity of the network.

*Hanoch Levy is on a sabbatical from the School of Computer Science, Tel Aviv University, Israel.

[1] In [11], the safe operating area of TCP's congestion control algorithm is discussed in terms of packet error rates and bandwidth×delay product.

AUTONOMICS 2007, 28-30 October 2007, Rome, Italy

DOI 10.4108/ICST.AUTONOMICS2007.2225

In simulation experiments [8] with a simple hop-by-hop transport protocol in a mobile wireless scenario, we observed a substantial gain over TCP if the connectivity of the network is disrupted frequently. However, these results do not lend themselves to generalization. In [7] we used a Bernoulli loss model and showed that under that model the hop-by-hop scheme is superior to the end-to-end scheme.

In an intermittent-connectivity environment, disruption periods of a connection might be lengthy and therefore a loss model that accounts for such duration is called for. We therefore choose to generalize the Bernoulli loss model used in [7] and model the packet losses via an on/off loss process, capturing extensive periods of disruption. Under this loss model we study the behavior of a single source–destination pair connected over multiple hops.

We analytically evaluate the end-to-end and hop-by-hop retransmission mechanisms. In the end-to-end case, lost packets are retransmitted by the source and in the hop-by-hop case, the retransmission is initiated by the node prior to the hop where the loss occurred. Our first set of results somewhat contradicts the Bernoulli loss model results and might be viewed, at first glance, as surprising. The finding is that the end-to-end scheme can be superior to the hop-by-hop scheme; this is in contrast to the Bernoulli loss findings where hop-by-hop is always superior. The superiority is in terms of higher delivery ratio (probability of successful delivery of a packet). More specifically, we find that under relatively long disruption periods, end-to-end tends to perform better, while under relatively short disruption periods it performs worse than hop-by-hop transport.

Examination of these findings reveals that the end-to-end superiority results from the fact that its natural retransmission interval (in the cases of packet loss) is significantly larger than that of the hop-by-hop scheme. This "patience" turns out to pay off in a network with long duration failures.

Having made this observation we introduce a variation of the basic hop-by-hop strategy, which we call *spaced hop-by-hop* and which deliberately uses a longer retransmission interval. We then analyze this scheme and compare it to both the end-to-end and the basic hop-by-hop scheme.

Our numerical results show that spaced hop-by-hop is superior, in terms of delivery ratio, to the end-to-end scheme for all durations of the disruption periods, and over a wide set of parameters. It is also superior to the basic hop-by-hop scheme for medium and long disruption durations.

Having established these findings for constant (over time) retransmission intervals, we then examine the spaced hop-by-hop scheme under time-varying exponential retransmission intervals. Such a scheme can be useful in driving the spaced hop-by-hop protocol into the appropriate retransmission interval. Needless to say that such dynamic exponential strategy has been widely used, e.g., in TCP. The relative comparison of the policies under the exponential retransmission yields results similar to those derived for the constant retransmission interval.

In prior work as far back as 1976, the two fundamental transport control principles have been compared, but under different assumptions. Gitman [6] evaluates delay and utilization under hop-by-hop and end-to-end retransmission in a similar analytical model of an early wireless network, but with uncorrelated packet loss and unlimited number of transmissions. He observes that hop-by-hop retransmission leads to lower delay and higher utilization for multi-hop paths and/or lossy links. In [2], DeSimone et al. investigate the crucial interaction between link-layer and transport-layer retransmissions analytically; they observe a positive impact of link-layer retransmissions on end-to-end throughput only if the link loss rate exceeds a threshold.

The rest of this paper is organized as follows. In the next section, we derive bounds for hop-by-hop and end-to-end transport and review our previous work based on simulation and analysis. In Sec. 3, we present our simple analytical connection model; numerical results follow in Sec. 4. In Sec. 5, we discuss the implications of our results and outline future work. In Sec. 6, we conclude the paper.

2. MOTIVATION

Hop-by-hop transport is intuitively expected to exhibit some performance advantages over end-to-end transport, as it recovers loss locally. For a chain topology scenario, it is also straightforward to show analytically that the upper bound on the achievable throughput is higher with hop-by-hop transport. Consider a source–destination pair communicating over a chain of H wireless links (hops). Assume that every link is only available a fraction q of the time, as a result of the wireless environment and node mobility. An upper bound for the achievable link throughput can be obtained under the following simplifying assumptions.

With end-to-end transport control, data can only be transferred when *all* links are available contemporaneously. During those periods the throughput T^e is bounded by the link capacity, C. Given the link loss rate q, we can write $T^e \leq q^H \cdot C$. In contrast, hop-by-hop transport is capable of utilizing individual links given that the node at the sending end of each hop has data to send. Thus, the throughput of hop-by-hop transport T^h is bounded by the fraction of hops that are available over the path length, so that the upper bound is given by $T^h \leq q \cdot C$. For all values of $0 < q < 1$, it holds that $T^h \geq T^e$.

2.1 A simulation case study: TCP vs. SAFT

To derive a first insight to the achievable gain of hop-by-hop transport in intermittent-connectivity scenarios, we performed simulation experiments. The network scenario involved sparse mobile ad hoc networks; the combination of mobility with sparse population resulted in frequent route disruptions and intermittent end-to-end connectivity. We compared the performance of TCP NewReno [4] to that of our own hop-by-hop protocol implementation called SAFT (Store-And-Forward Transport). SAFT provides a service interface identical to that of TCP, but transfers data hop-by-hop; the protocol is described in [8].

In our simulation study, 30 nodes moved according to the random waypoint mobility model within an area of 1000m $\times$ 3000m. Since the wireless range is only 250m, disruptions were very likely. Routes were established on-demand via the AODV [17] routing protocol. Our application scenario was simple. At the beginning of the experiment, 10 source–destination pairs were chosen randomly among all nodes. The source nodes started transmitting 10 messages of 100KB each during the first 100 seconds. Hence, the total amount of data to be transferred by the network was 10'000KB.

In Fig. 1, we plot the data transfer progress over time for 10 connections from a random simulation run. The bold curve indicates the total amount of data transferred as a percentage of 10'000KB. With TCP, the complete transfer took one hour; SAFT delivered the same amount of data within one quarter of the time. The plotted scenario illustrates that the progress of the data transfer with SAFT was much steadier than with TCP. The plot of the TCP progress shows periods where the destination nodes did not receive any data. This alternation between transfer and idle periods is also visible in the hop-by-hop protocol's plot, but the idle periods are shorter.

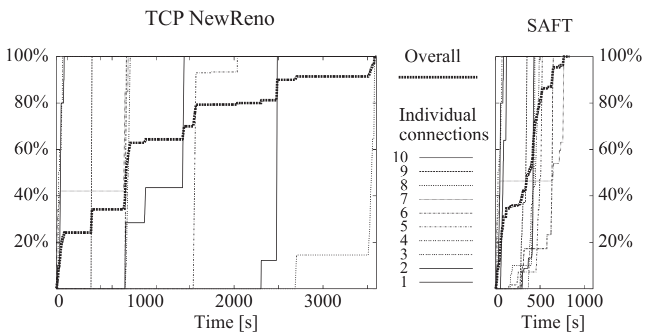


Figure 1: **Transfer progress**. 100% ≙ 10'000kB

Our analysis of the trace files revealed that TCP transmitted data primarily over single-hop connections, and typically, only one connection was transmitting at a time. In contrast, SAFT used a mix of single- and multi-hop routes, sharing the limited bandwidth more effectively among multiple connections. Additionally, the hop-by-hop protocol started transmitting data much earlier than TCP. The late start of TCP is partly due to the frequent pseudo route failures that occurred with this protocol. Such failures were not due to node mobility but rather caused by TCP's bandwidth probing mechanism.

Overall, the simulation experiments suggested that, in the considered scenarios, our hop-by-hop protocol is able to make better use of the communication opportunities. However, the performance of the transport protocol is dependent on many variables (transport layer configuration, routing protocol, mobility), so that the simulation results can only be viewed as a case study. To enable more positive statements about the relative performance of the two transport alternatives, we turned our focus on analytical modeling.

2.2 End-to-end vs. hop-by-hop transport under Bernoulli loss

In [7] we carried out a first comparison of the two transport alternatives under simple models for the network and the link loss process. We assumed a chain topology with H hops and a single source–destination pair. At every hop i, the packet loss process was Bernoulli with parameter q_i; in case of transmission failure, the packet was retransmitted up to $L-1$ times. With end-to-end transport, retransmission was carried out at the source; in the hop-by-hop case, the packet was retransmitted locally over the hop where the failure occurred. We computed the probability that a packet arrives at the destination and the expected number of required link transmissions as a function of H and L.

In Fig. 2, we show these two metrics for the 5-hop chain, with $q_i = q, \forall i, 1 \leq i \leq H$, and a maximum of $L = 7$ and $L = 15$ transmissions. Fig. 2a suggests that uncorrelated loss of only a few percent can be tackled well by both transport control principles. As the link loss probability q increases, both schemes degrade gracefully, the delivery ratio of the end-to-end scheme dropping more rapidly. As expected, increasing the number of transmissions results in increase of the delivery ratio under high loss probabilities. The hop-by-hop scheme is superior in terms of delivery ratio to the end-to-end scheme at both settings of L for all loss probabilities q. For most q values this advantage of hop-by-hop is combined with fewer link transmissions, as shown in Fig. 2b. Only at high q values, beyond $q = 0.5$ in these plots, are the transmissions with hop-by-hop more but this is the penalty for achieving a non-zero packet delivery probability where the end-to-end scheme completely collapses.

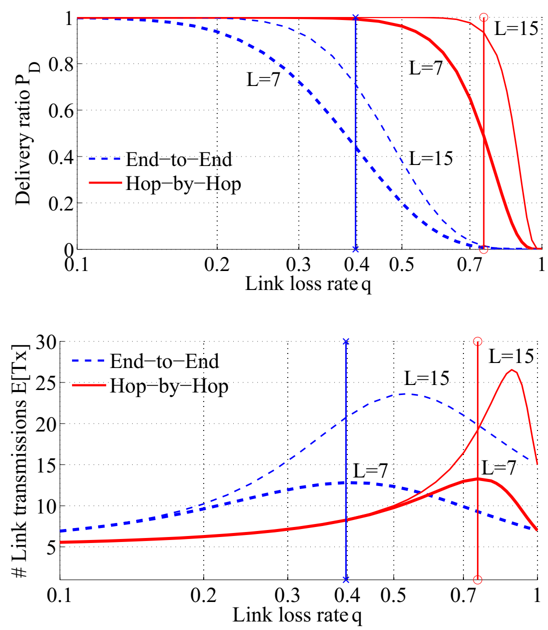


Figure 2: **Bernoulli loss process;** 5 hops

To summarize, under Bernoulli packet loss, hop-by-hop transport is more efficient in the use of link transmissions achieving higher delivery probabilities for any loss probability $q \in [0, 1)$. Our evaluation confirms the well-known fact that hop-by-hop retransmission is more effective against uncorrelated loss than the end-to-end alternative. However, under intermittent connectivity, packet loss is correlated. In the rest of the paper, we study the relative performance of the protocols under a model that is more appropriate for intermittent-connectivity scenarios and taking into account real-world transport protocol retransmission policies.

3. MODELING AND ANALYSIS

In this section, we introduce our model of end-to-end and hop-by-hop transport over an intermittently-connected network path.

3.1 Modeling intermittent connectivity

We still consider a chain topology where the source and destination communicate over a finite number of wireless hops, H. The link state in each hop is modeled by a two-state discrete-time Markov chain (DTMC) [5] alternating between two states: a "good" or "on" state, during which packet transmissions are successful and a "bad" or "off" state, during which communication between the two hop end points is lost. Transitions from the good (bad) to bad (good) state for each hop i occur at discrete time steps, equal to the packet transmission time, with probability $p_{gb}(i)$ and $p_{bg}(i)$, respectively; the corresponding complementary probabilities, $p_{gg}(i) := 1 - p_{gb}(i)$, $p_{bb}(i) := 1 - p_{bg}(i)$, refer to the probability of staying at the good (bad) state, as depicted in Fig. 3.

Compared with its Bernoulli counterpart, this loss model can capture correlations in the loss process, which are common over wireless links. Moreover, it can be viewed as a simple model for link disruptions.

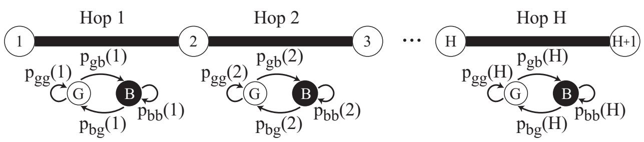


Figure 3: **On/off loss model.**

In the context of this paper, we ignore dependencies between successive hops; the link state in each hop varies independently from that of other hops. This would be the case, for example, when transmissions over successive hops occur over different radio channels. It is well known that for the described two-state DTMC, the expected durations of the two states are given by

$$E_B^i = \frac{1}{p_{bg}(i)}, \qquad E_G^i = \frac{1}{p_{gb}(i)}; \tag{1}$$

the steady-state probabilities of the good and bad state are given by

$$P_B^i = \frac{p_{gb}(i)}{p_{bg}(i) + p_{gb}(i)}, \qquad P_G^i = \frac{p_{bg}(i)}{p_{bg}(i) + p_{gb}(i)}; \tag{2}$$

and the mean link loss rate q^i over the i^{th} hop is defined as

$$q^i := P_B^i. \tag{3}$$

Moreover, the link state transition probability matrix for hop i is:

$$\mathbf{P_i} = \begin{pmatrix} p_{bb}(i) & p_{bg}(i) \\ p_{gb}(i) & p_{gg}(i) \end{pmatrix} \tag{4}$$

and the k-step transition probabilities are the elements of matrix $\mathbf{P_i^{(k)}}$, defined as

$$\mathbf{P_i^{(k)}} := \begin{pmatrix} p_{bb}^{(k)}(i) & p_{bg}^{(k)}(i) \\ p_{gb}^{(k)}(i) & p_{gg}^{(k)}(i) \end{pmatrix} := (\mathbf{P_i})^{\mathbf{k}} \tag{5}$$

with elements $p_{gb}^{(k)}(i)$ and $p_{bg}^{(k)}(i), 1 \le i \le H$, denoting the probabilities of switching from good(bad) to bad(good) state within k steps of the DTMC evolution, respectively.

3.1.1 Loss correlation and disruption duration

We say that the process represented by the transition probability matrix $\mathbf{P_i}$ is *positively correlated* if $P_B^i < p_{bb}(i) = 1 - p_{bg}(i) \Leftrightarrow p_{bg}(i) + p_{gb}(i) < 1$. This means that the conditional probability of the next slot to be "off" (conditioned on the current slot being "off") is higher than the steady state probability of being in "off" state.

The process correlation is zero when the probability of entering the on or the off state is independent of the current state. This holds when $p_{bg}(i) = p_{gg}(i) \Leftrightarrow p_{gb}(i) = p_{bb}(i)$, which is equivalent to $p_{bg}(i) + p_{gb}(i) = 1$. Using the definitions of q and $E[B]$ from (2) and (3), the disruption duration $E[B]$ where the on/off loss model has zero correlation and thus behaves like the Bernoulli model is $E[B] = 1/(1-q)$.

Finally, the process is called *negatively correlated*, when $E[B]$ is less than $1/(1-q)$; this means that if a link is in the off state, it is more likely to be in the on state in the next time slot, i.e., $p_{bg}(i) > p_{bb}(i)$, and vice-versa for a link that is in the on state.

3.2 Transport protocol retransmission policy

The number of transmissions is finite and limited to L for both the end-to-end and hop-by-hop transport schemes. In the end-to-end scheme, this is the number of attempts made by the source node; for the hop-by-hop scheme, this is the number of attempts made by the source and each of the intermediate nodes. Hence, the maximum number of per packet link transmissions for both schemes is upper bounded by the same number $H \cdot L$. Note that in the end-to-end case, this number is reached if the packet traverses the first $H-1$ hops successfully and then fails at the H^{th} hop; whereas, under the hop-by-hop scheme, this number of transmissions is expended, if at every hop the packet fails in the first $L-1$ transmission attempts but succeeds in the L^{th} transmission attempt.

Most real-world network protocols involving data retransmission employ some sort of back-off algorithm, usually exponential, for spacing retransmissions in time. The end-to-end and the hop-by-hop transport protocol we present in this section may follow *any* retransmission policy. We use $rto(k)$ to denote the retransmission timeout (RTO) between the $(k)^{th}$ and the $(k+1)^{th}$ transmission attempt, $1 \le k \le L-1$, i.e., after the k^{th} unsuccessful transmission, the retransmitting node waits for $rto(k)$ time slots before initiating the $(k+1)^{th}$ transmission.

For both transport alternatives, hereafter noted with superscripts e and h, we derive two metrics for the H-hop chain topology; the probability of successful packet delivery probability $P_D^{e(h)}$, and the expected number of link transmission attempts $E[Tx^{e(h)}]$ for each packet. Note that for $L = 1$, there is only one transmission attempt and no retransmissions. Thus, both schemes behave the same. The probability of successful delivery is given by $P_D = \prod_{i=1}^{H} P_G^i$ and the expected number of transmissions is $E[Tx] = \sum_{i=1}^{H-1} i \cdot (1 - P_G^i) \prod_{k=1}^{i-1} P_G^k + H \cdot \prod_{k=1}^{H-1} P_G^k$. In the remainder of this section, we assume that there is at least one retransmission allowed, i.e., $L \ge 2$.

3.3 Analysis of end-to-end transport

We use the framework of absorbing Markov processes [5] to model the end-to-end transport behavior on/off loss process; we give a brief introduction to this subject in the next paragraphs.

3.3.1 Absorbing Markov processes

Absorbing Markov processes are Markov processes that will eventually stop at one or more absorbing states. The transition probabilities $p(i, j)$ from an absorbing state i towards all other states $j, j \ne i$ will be zero, whereas they equal 1 when $i = j$. The remaining non-absorbing states of the process are called transient states. For an absorbing Markov process with n absorbing states and $N-n$ transient states, the probability transition matrix $\mathbf{X}$ can be written as the concatenation of four different matrices

$$\mathbf{X} = \begin{pmatrix} T & R \\ 0 & I \end{pmatrix}$$

In the matrix $\mathbf{X}$, the sub-matrix $\mathbf{T}$ is an $(N-n)\times(N-n)$ matrix, with $T(i,j)$ being the transition probabilities among transient states, whereas the sub-matrix $\mathbf{R}$ is the $(N-n)\times n$ matrix with the transition probabilities from transient states towards absorbing states. The sub-matrix $\mathbf{I}$ in $\mathbf{X}$ is the $n\times n$ identity matrix of size n, corresponding to the transition probabilities amongst absorbing states, and the sub-matrix $\mathbf{0}$ is the all-zero $n \times (N-n)$ matrix with the transition probabilities from absorbing states to transient states.

The description of the absorbing Markov process is complemented by the $1 \times (N-n)$ initial probability vector $\mathbf{e}$, whose entries $e(i)$ express the probabilities that the process starts at state i.

With these matrices at hand, it is possible to define the *fundamental* matrix $\mathbf{Q}$ of the absorbing process

$$\mathbf{Q} = (\mathbf{I} - \mathbf{T})^{-1} \tag{6}$$

which is most helpful in the computation of several process metrics. For example, the entry $Q(i,j)$ denotes the expected number of visits to state j, when starting from state i, before ending up in an absorbing state. Likewise, the $U(i,j)$ element of the product matrix $\mathbf{U} = \mathbf{Q}\cdot\mathbf{R}$ is the conditional probability of ending in absorbing state j, given that the process was initiated at state i. Finally, if the number of steps before the process reaches one of the absorbing states and terminates is denoted by K, then its expected value is given by

$$E[K] = \mathbf{e}\cdot\mathbf{Q}\cdot\mathbf{1} \tag{7}$$

where $\mathbf{1}$ is the $1 \times (N-n)$ all-one vector.

3.3.2 End-to-end transport as absorbing Markov process

The state space of the Markov process in our case is finite and features two absorbing states; they correspond to packet delivery success and failure, respectively. Each process state combines the link states $\{s_i\}$ of the individual hops in the chain shown in Fig. 3, $s_i \in \{G, B\}, 1 \le i \le H$, the hop identity $h, 1 \le h \le H$, and the transmission attempt $l, 1 \le l \le L$. The process (packet transmission) starts from one of the 2^H different states with $h = l = 1$ and progresses towards higher h and/or l, till it reaches one of the two absorbing states. The number of non-absorbing states is 2^H for $h = 1$, and 2^{H-1} for $h > 1$, yielding a total of $W = 2^H \cdot (1 + (H-1)/2) \cdot L$ states.

The application of equations (6)-(7) requires the derivation of matrices T and R. The process progresses from states with $h = m, 1 \le m \le H$, and $l = k, 1 \le k \le L$ to states with $h > 1, l = k$, in case of success, or states $h = 1, l = k+1, k \le L-1$, in case of failure. In the second case, when the failure takes place over hop h, the transition probabilities involve the $(2H-h)$-step transition probabilities of the link state processes $\{s_i\}$. The "failure" absorbing state is reachable from states with $l = L, h = m, 1 \le m \le H$, with $p_{gb}(k)$ or $1 - p_{bg}(k)$, depending on the actual state, whereas the "success" state is reached from states featuring $h = H$, with probabilities $p_{gg}(H)$ or $p_{bg}(H)$, again depending on the actual state. An example of the Markov absorbing process for $H = 2, L = 2, rto(k) = 2H\ \forall k$ is shown in Fig. 4, which also depicts all possible transitions amongst the process states as well as values of the transition probabilities.

Combining the aforementioned probabilities into matrices $\mathbf{T}$ and $\mathbf{R}$, the probability of successful delivery, P_D^e, is the second element of the 1×2 matrix $\mathbf{e}\cdot\mathbf{U}$, whereas the expected number of transmissions, $E[Tx^e]$, is given by((7)), where $\mathbf{e}$ is the $1 \times W$ initial state probability vector.

3.4 Analysis of hop-by-hop transport

We assume that all hops are independent; therefore, we can model hop-by-hop transport based on mean-value analysis. When packets are acknowledged and retransmitted on a hop-by-hop basis, the total number of link transmissions Tx^h, required for end-to-end packet delivery under maximum L transmissions per hop can be expressed as the sum of the transmissions in each one of the H hops

$$Tx^h \quad = \quad M_{1,L}^h + M_{2,L}^h + \ldots + M_{H,L}^h.$$

where $1 \le M_{1,L}^h \le L$ and $0 \le M_{i,L}^h \le L, 2 \le i \le H$. Taking mean values we have

$$E[Tx^h] \quad = \quad E[M_{1,L}^h] + E[M_{2,L}^h] + \ldots + E[M_{H,L}^h].$$

The expected number of link transmissions over hop i is conditioned on the probability that the packet makes it to hop i, i.e., it successfully crosses all prior $i-1$ hops.

$$E[M_{i,L}^h] \quad = \quad E[M_{i,L}^h]_c \cdot P(\textit{packet crossed all prior } i-1 \textit{ hops}).$$

where $E[M_{i,L}^h]_c$ denotes the expected number of link transmissions over hop i, under the condition that the packet has traversed all $i-1$ previous hops.

The probability of success for the first transmission over hop i is simply P_G^i; the probability of requiring two attempts to cross hop i is $P_B^i \cdot p_{bg}^{(rto(1))}(i)$ and, generalizing, that of requiring n attempts

$$\begin{aligned} P_{i,L}^n &= P_B^i \cdot \prod\nolimits_{l=1}^{n-2} p_{bb}^{(rto(l))}(i) \cdot p_{bg}^{(rto(n-1))}(i), \; n \ge 2 \\ &= (1 - \sum_{l=1}^{n-1} P_{i,L}^{l-1}) \cdot p_{bg}^{(rto(n-1))}(i), \; n \ge 2, \end{aligned} \tag{8}$$

where $P_{i,L}^1 := P_G^i$; the definition of the k-step transition probabilities $p_{bb}^{(rto(l))}(i) := p_{bb}^{(k)}(i)\,|_{k=rto(l)}$, $p_{bg}^{(rto(n-1))}(i) := p_{bg}^{(k)}(i)\,|_{k=rto(n-1)}$ is given in (5).

The packet manages to cross the i^{th} hop within L link transmissions with probability

$$P_{i,L}^{succ} = \sum_{k=1}^{L} P_{i,L}^k \tag{9}$$

and the expected number of link transmissions over hop i, conditioned upon the successful packet transmission up to hop $i-1$ equals

$$E[M_{i,L}^h]_c = \sum_{k=1}^{L-1} k \cdot P_{i,L}^k + L \cdot \sum_{j=1}^{L-1} (1 - P_{i,L}^j).$$

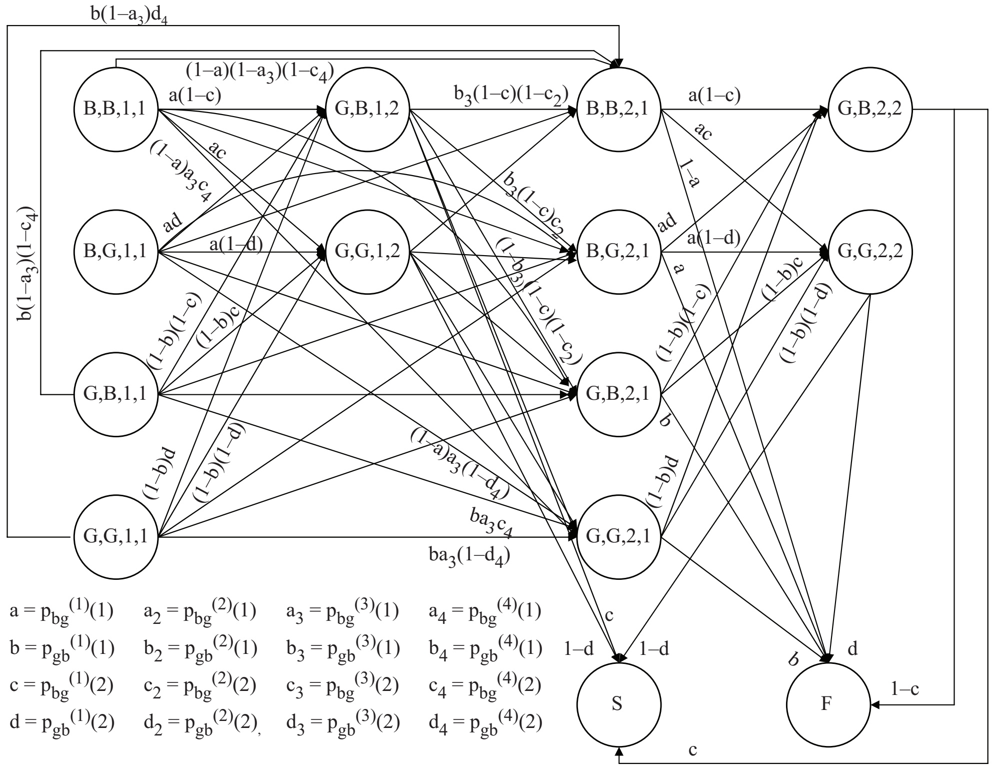


Figure 4: Example of absorbing Markov chain with exemplary transition probabilities; $H = 2,\ L = 2,$ and $rto(k) = 2H\ \forall k$.

Therefore (8) yields for the expected number of link transmissions

$$E[Tx^h] = \quad E[M^h_{1,L}] + \sum_{k=2}^{H} E[M^h_{k,L}]_c \cdot \prod_{j=1}^{k-1} P^{succ}_{j,L}. \qquad (10)$$

Finally, the probability of successful delivery of the packet becomes

$$P^h_D = \prod_{i=1}^{H} P^{succ}_{i,L}. \qquad (11)$$

In the next section we present numerical results derived from our model of end-to-end and hop-by-hop transport under intermittent connectivity.

4. NUMERICAL RESULTS

In this section, we use numerical results to compare the performance of end-to-end and hop-by-hop transport. We consider the probability of successful delivery to the destination, P_D, (henceforth referred to as delivery ratio) and the expected number of link transmissions $E[Tx]$, where Tx is the number of link layer transmissions incurred by the protocol in the process of sending a packet until the packet either reaches the destination or it is discarded after L failed attempts. Note that we discuss the implications of the numerical results and the limitations of our model in Sec. 5.

In order to investigate the effect of various disruption durations, we evaluate end-to-end and hop-by-hop transport over a range of mean disruption durations $1 \leq E[B] \leq 1000$. We consider both a low loss rate of $q = 5\%$ as a reference point for the situation in a rather well-connected network as well as a value of $q = 20\%$. The length of the chain topology is $H = 5$ hops and the number of transmission attempts is limited by $L = 5$.

For the following discussion, we introduce a measure that we call *transmission period*, T_{Tx}. The transmission period denotes the maximal period of time along which a packet is retransmitted before it is discarded; thus if any hop is in the off state for longer than this period, the protocol will fail to deliver the packet. The transmission period is determined by the maximum number of transmissions L and by the retransmission timeout algorithm of the protocol.

4.1 End-to-end vs. hop-by-hop transport

We compare end-to-end vs. hop-by-hop transport with the following parameters. The retransmission timeout (RTO) of the end-to-end scheme is set to the minimum value of one round-trip time: $rto(k) := 2H, k = 1, 2, \ldots, L-1$; the timeout of the hop-by-hop transport protocol is set to the minimum of one time slot, i.e., $rto(k)^h := 1, k = 1, 2, \ldots, L-1$. Such a short retransmission timeout seems feasible if the transport protocol relies on the acknowledgment scheme of the link layer to determine losses. Under these settings, the transmission period of the end-to-end scheme is $T^e_{Tx} = 2H \cdot L = 50$ and for the hop-by-hop scheme it is $T^h_{Tx} = L = 5$ time slots.

In Fig. 5, we plot the delivery ratio and the expected number of link transmissions as functions of the mean disruption duration. The mean link lifetime $E[G]$ changes accordingly to achieve the given link loss rates of $q = 5\%$ and $q = 20\%$.

In the upper plot of Fig. 5, we show the delivery ratio, and we can see three trends that are common to both values of q. First, at the very low end of expected disruption

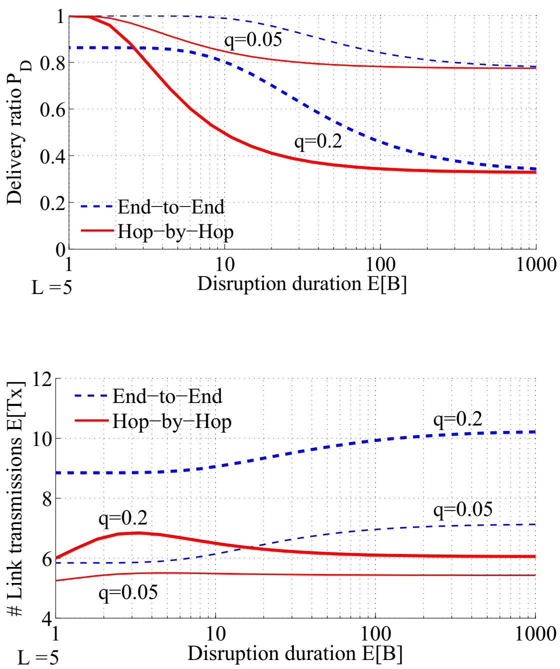


Figure 5: **End-to-end vs. hop-by-hop transport performance** as a function of the mean disruption duration $E[B]$; constant link loss rates $q = 0.05$ and $q = 0.2$

duration values, the hop-by-hop protocol achieves a delivery ratio very close to unity, even at $q = 0.2$. In contrast, the end-to-end scheme is limited by an upper bound; even for the shortest disruption duration $E[B] = 1$. For $q = 0.05$, the end-to-end scheme achieves a delivery ratio of $P_D = 99.94\%$; for $q = 0.2$ it delivers 86% of the packets.

Second, the delivery ratio of the hop-by-hop transport protocol is much more sensitive to the mean disruption duration and thus becomes significantly inferior at medium and long disruption durations.

Third, for adequately high values of $E[B]$, both schemes reach an identical performance floor, which corresponds to what can be achieved under the Bernoulli loss model without retransmissions, i.e., if $L = 1$.

As a result of these trends, there is a crossover of the delivery ratio values of the end-to-end and the hop-by-hop transport protocol. Note that in the plot, this crossover can only be seen for $q = 0.2$, where it occurs around $E[B] = 2.7$.

Concerning the number of link transmissions given in the bottom plot of Fig. 5, the hop-by-hop scheme results in a lower number of link transmissions for both values of q. It has a maximum at a certain value of the disruption duration that could be characterized by a sharp decrease in the delivery ratio. The end-to-end protocol spends an increasing number of link transmissions as the disruption durations become longer; even at a value of $E[B]$ where the delivery ratio is close to the performance floor, $E[Tx^e]$ is still increasing.

Hop-by-hop transport under short disruption durations. The higher delivery ratio of the hop-by-hop protocol for short disruption periods can be explained analytically. Irrespective of the number of transmissions, the delivery ratio of the hop-by-hop scheme always begins at the same value for $E[B] = 1$; $E[B] = 1$ implies that $p_{bg} = 1$, i.e., every time the link enters the off state, it will stay there for one time slot. Therefore, with $rto(k) = 1,\ k = 1, 2, \ldots, L-1$, for every $L \geq 2$, the probabilities $P^n_{i,L}$ in (8) equal P_G for $n = 1$, P_B for $n = 2$, and 0 for $n > 2$, the probabilities $P^{succ}_{i,L}$ in (9) become 1 and the number of expected link transmissions in (10) reduces to

$$E[Tx^h] \quad = \quad \sum_{i=1}^{H} E[M^h_{i,L}] = \sum_{i=1}^{H} (P^i_G + 2 \cdot (1 - P^i_G))$$

which is independent of L for short disruptions.

However, if $E[B]$ is longer than the transmission period T^h_{Tx}, the hop-by-hop protocol collapses because a necessary condition for successful delivery is that no hop be disrupted longer than T^h_{Tx}.

End-to-end transport under long disruption durations. Since the end-to-end scheme covers a longer period with its transmission attempts, it is more likely to be still transmitting when a disrupted hop has switched to the on state. We investigated the performance of both schemes also with only three instead of five hops and found that the delivery ratio is higher for short mean disruption durations but drops at considerably lower values of $E[B]$. This is due to the definition of the retransmission timeout of the end-to-end scheme, which depends on the number of hops; and with three hops, the transmission period covers only $2H \cdot L = 30$ instead of 50 slots.

Performance floor of delivery ratio. The performance floor that both protocols approach as the mean disruption duration increases is a result of the strong positive correlation in the loss process for high values of $E[B]$. Since the probability of switching from the off to the on state is the inverse of the mean duration of the off state, i.e., $p_{bg} = 1/E[B]$, this probability approaches zero as $E[B]$ increases. The same holds for p_{gb} and $E[G]$. For large values of $E[B]$, the loss rate q, which is also the probability of finding the hop in the off state, determines success or failure of a transmission, independent of when this transmission occurs. The limited number of transmission attempts are likely to all occur during a period where the links do not switch from one to the other state; thus either the first transmission succeeds, or none of the retransmissions will succeed either. Therefore, at high values of $E[B]$, the delivery ratio is the same with both protocols and is equal to that under Bernoulli loss without retransmissions.

Crossover of end-to-end and hop-by-hop transport. The crossover between the curves of the two schemes (for $q = 0.2$ at $E[B] = 2.7$) highlights the two physical properties governing the relative performance of these schemes. The first is the work conservation of hop-by-hop, which crosses each hop only once; this property is not shared by end-to-end and thus hop-by-hop enjoys a relative advantage. The second is the relative length of the transmission period and the disruption period. Having the transmission period longer than the disruption period is an advantage since if the transmission period is too short, packet retransmissions are likely to fail. As the hop-by-hop scheme we employ sends retransmitted packets in consecutive time slots, it is susceptible to lengthy disruption periods. The end-to-end scheme spaces retransmissions over time because it requires a longer retransmission timeout and thus has a relative advantage in this range.

The susceptibility of hop-by-hop transport to the disruption duration can be amended by increasing its transmission period. One way to achieve this would be to increase the transmission limit L; however, at high values of q, this change would also result in a much higher number of link transmissions. As a more efficient alternative, we investigate in the next subsection a hop-by-hop scheme that deliberately spaces retransmissions over time, aiming to cover a larger period transmitting with the same number of transmission attempts. We expect that such a scheme should be better equipped to perform well in intermittent-connectivity scenarios where we expect periods of disruption to be extensive.

4.2 Spaced hop-by-hop transport

We compare with the two previously studied schemes a protocol that is the same as the hop-by-hop scheme except that we introduce an idle period between successive transmission attempts by increasing the retransmission timeout period. In the following, we refer to this scheme as *spaced hop-by-hop transport* and denote quantities relating to it by the superscript s. When referring to the previously introduced hop-by-hop protocol, we may add the qualifier basic for clarity. For the spaced hop-by-hop protocol, we choose the retransmission timeout to be considerably longer than the minimum value of $T^h_{RTO} = 1$ we use in the basic hop-by-hop scheme and set it equal to that of the end-to-end protocol, i.e., $T^s_{RTO} = T_{RTT} = 2H$ time slots. Under this setting, the transmission period of spaced hop-by-hop transport is equal to that of the end-to-end scheme, i.e., $T^s_{Tx} = 2H \cdot L = 50$ time slots.

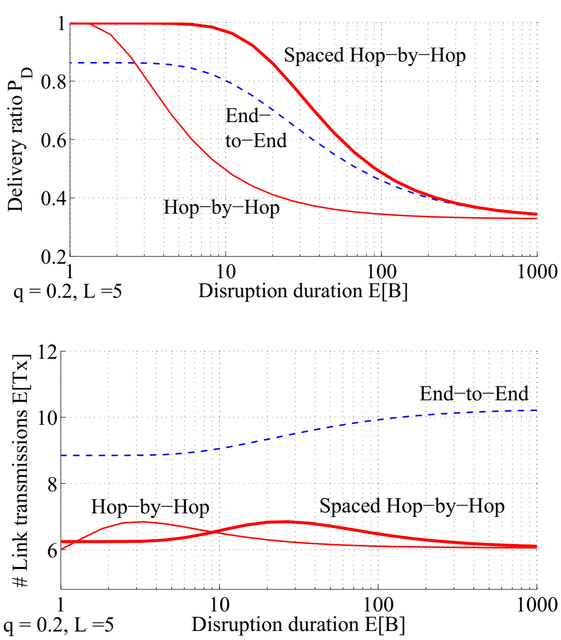


Figure 6: **Spaced hop-by-hop vs. hop-by-hop vs. end-to-end transport.** $q = 0.2$

In Fig. 6 we plot the delivery ratio and the expected number of link transmissions as a function of the disruption duration, for $q = 0.2$. The delivery ratio of spaced hop-by-hop at $E[B] = 1$ is slightly lower than that of basic hop-by-hop, but it is much less sensitive to the mean disruption duration, resulting in a substantially higher delivery ratio for medium and long disruption durations. Compared to the end-to-end protocol, we see that the spaced hop-by-hop protocol achieves a higher delivery ratio for all values of the mean disruption duration. As $E[B]$ increases, all three schemes approach the performance floor discussed in the previous subsection.

Regarding the number of link transmissions spent per packet $E[Tx]$, both hop-by-hop protocols spend about the same number of transmission and the maximum value of the spaced hop-by-hop scheme is located at a higher value of $E[B]$. For all values of $E[B]$, $E[Tx]$ is lower with both hop-by-hop protocols than with the end-to-end protocol.

We also evaluate the spaced hop-by-hop protocol at $q = 0.05$, but in order to highlight differences among the schemes, we allow at most one retransmission. In Fig. 7 we plot the delivery ratio as a function of the disruption duration, for $q = 0.05$ with a maximum of $L = 2$ transmissions. Note that in this plot the range on the vertical axis is not $[0, 1]$ but $[0.75, 1]$. The basic hop-by-hop scheme reaches a delivery ratio close to one at a disruption duration of $E[B] = 1$. At this value of $E[B]$, the spaced hop-by-hop protocol reaches only a maximum delivery ratio of 98.8% and the end-to-end protocol is limited to 94.9%. As $E[B]$ increases, the delivery ratio of the basic hop-by-hop protocol immediately declines whereas the ratio of spaced hop-by-hop and end-to-end only drop around $E[B] = 4$. The spaced hop-by-hop protocol provides a higher delivery ratio than both the end-to-end and the basic hop-by-hop protocol for the full range of mean disruption durations we evaluate.

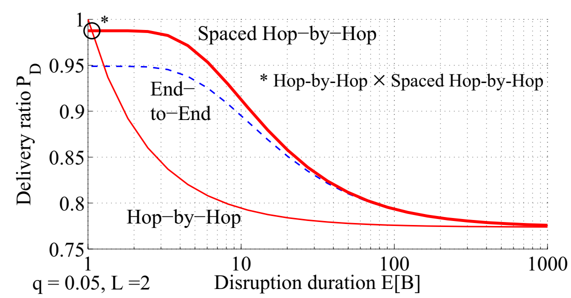
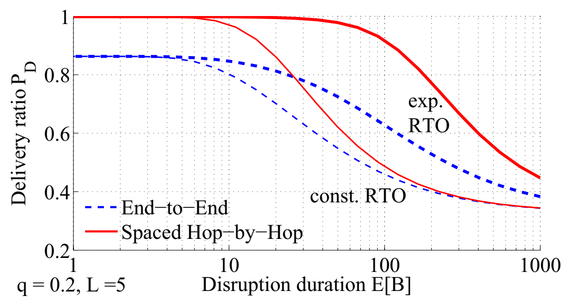


Figure 7: **Spaced hop-by-hop vs. hop-by-hop vs. end-to-end transport.** $q = 0.05, \mathbf{L = 2}$

Basic vs. spaced hop-by-hop at short disruption durations. As shown analytically in Sec. 4.1, the high delivery ratio of the basic hop-by-hop protocol for short disruption periods is due to its retransmission policy that sends packets back-to-back. The relative performance between the basic hop-by-hop protocol and the spaced hop-by-hop protocol changes at the point where the correlation of the loss process switches from negative to positive, i.e., where the correlation is zero and $p_{gb} + p_{bg} = 1$ (cf. Sec. 3.1.1). For all greater values of $E[B]$, spaced hop-by-hop provides a higher delivery ratio than basic hop-by-hop. Under the on/off loss model, we can define analytically the range of disruption durations where it pays off to retransmit packets back-to-back (basic hop-by-hop) and where it is more beneficial to space transmission attempts over time (spaced hop-by-hop). For $q = 0.05$, the on/off loss process is negatively correlated for values of $E[B] < 1/(1 - q) = 1/0.95 = 1.05$; and within this range, basic hop-by-hop dominates spaced hop-by-hop.



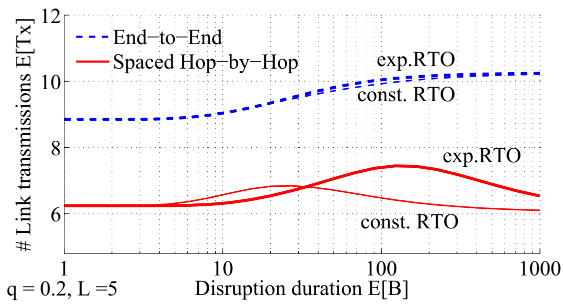


Figure 8: **Exponential vs. constant retransmission timeout.** $q = 0.2$

For larger values of $E[B]$, the correlation is positive and it pays off to space transmission attempts over time; thus the spaced hop-by-hop protocol dominates the basic variant for all values of $E[B] > 1/(1-q)$.

Spaced hop-by-hop vs. end-to-end transport. If the retransmission timeout of the spaced hop-by-hop protocol is equal to that of the end-to-end protocol, we observe that spaced hop-by-hop dominates the end-to-end scheme in delivery ratio and expected number of link transmissions for both settings of q and for all values of $E[B]$ we consider.

We are currently attempting to establish these two dominance relationships fundamentally and we summarize our preliminary results in Sec. 5. The results of the spaced hop-by-hop protocol support our previous observation that the transmission period T_{Tx} determines to a large extent the response of a protocol to extensive periods of disruption. The spaced hop-by-hop and the end-to-end protocol can be tuned to a given disruption duration by adjusting L and $rto(k)$. However, in real-world mobile scenarios, the transmission limit cannot be arbitrarily high due to interference concerns and energy constraints. Moreover, the duration of disruptions is in general not known a priori. Therefore, most real-world implementations of retransmission protocols begin with a low retransmission timeout and then gradually increase it if the disruption persists. In the next section, we compare end-to-end and spaced hop-by-hop transport with constant and increasing retransmission timeout policies.

4.3 Exponential RTO

We compare the previously introduced end-to-end and spaced hop-by-hop scheme with both, a constant retransmission timeout $rto(k) := 2H, k = 1, 2, \ldots, L-1$ and an exponentially increasing timeout, defined by $rto(k) := T_{RTT} \cdot 2^{k-1} = 2H \cdot 2^{k-1}$. Thus, the exponential schemes have a timeout that starts at the same value of $rto(1) = 2H \cdot 2^0 = 10$ as the constant schemes, but then doubles for every retransmission, reaching a maximum of $rto(4) = 2H \cdot 2^3 = 80$ time slots between the fourth and fifth transmission attempt. Under this policy, the transmission period is given by $\sum_{k=1}^{L-1} 2H \cdot 2^{k-1} = 150$ time slots.

In Fig. 8, we plot the delivery ratio and the expected number of link transmissions for the end-to-end and the spaced hop-by-hop scheme under both a constant and an exponential retransmission timeout policy. We find that if both schemes employ the same retransmission policy, the spaced hop-by-hop scheme dominates end-to-end both in terms of delivery ratio and expected number of link transmissions over the whole range of $E[B]$ we consider. In the next section, we will draw conclusions from these results and discuss the limitations of the current model and outline future work.

5. DISCUSSION AND CONCLUSIONS

In the following, by *hop-by-hop transport*, we refer to the general concept of hop-by-hop transport; when we talk about a specific protocol implementation, we use the terms *basic* or *spaced hop-by-hop* protocol. We summarize below the main results coming out of the analysis in Sec. 3 and the numerical results presented in the previous section:

Delivery ratio The delivery ratio of the basic hop-by-hop tends to be higher than that of the end-to-end scheme for short disruption periods and lower for long periods. The performance of the schemes converges at infinite disruption duration, where the performance of both schemes resembles that under Bernoulli loss and no retransmissions. Considering spaced hop-by-hop we find that:

- The delivery ratio of spaced hop-by-hop is higher than that of end-to-end for all conditions studied.
- The delivery ratio of spaced hop-by-hop is higher than that of basic hop-by-hop in the region where the correlation of the loss process is positive (see Sec. 3.1.1 for the definition of positive correlation). It is lower when the correlation is negative.

Number of link transmissions The spaced hop-by-hop protocol always results in fewer link transmissions than the end-to-end scheme. Fewer retransmissions imply lower interference levels in the network and energy savings for the often power-constrained network nodes. In combination with the dominance in terms of delivery ratio, spaced hop-by-hop emerges as the most efficient transport alternative.

We are currently continuing the analytical work in this paper to draw definitive conclusions about the relative performance of the transport schemes we discussed. Preliminary results suggest that:

- If the spaced hop-by-hop and the end-to-end protocol use the same retransmission timeout policy, then there is **full dominance of spaced hop-by-hop over end-to-end transport** in terms of delivery ratio, expected number of link retransmissions, but also packet delivery latency, a metric that has not been looked into in this paper.
- For any spaced hop-by-hop policy with a retransmission timeout greater than one, there is **full dominance of spaced hop-by-hop over basic hop-by-hop transport** as long as the loss process is positively correlated, i.e., when $E[B] > 1/(1-q) \Leftrightarrow p_{bg} + p_{gb} < 1$

We intend to report on these results in a forthcoming paper.

5.1 Limitations of the model and future work

Our current models and analysis fail to capture the dynamics of the nodes on the path between the source and the destination. Node mobility and the operation of the underlying routing protocol may result in changes of the path length amidst the packet transfer. The absorbing Markov processes can be expanded to capture path length changes, although they cannot account for the change in the link states along the new path. We are currently working on implementing this extension.

A harder limitation of the current modeling framework is the assumption of independence in the link availability of successive hops (strictly speaking, the loss process in successive hops). This assumption may hold when different radio channels, spaced in time and/or frequency, are used over neighboring hops. However, in all real-world networks, interference between successive links results in substantial correlation amongst the transmission and loss processes in neighboring hops.

Finally, and relevant to the previous limitation, the work in this paper has considered a chain topology scenario with a single source–destination pair. Carrying out the comparison between the end-to-end and the hop-by-hop transport paradigm at the network-level, considering the distribution of nodes in space and the lengths of resulting paths, would reveal the impact of the transport scheme on system-level metrics such as the network capacity. We have already begun exploring this research direction, which necessitates a different modeling framework than the one presented in this paper.

6. SUMMARY

We investigated the performance of end-to-end vs. hop-by-hop transport under intermittent connectivity using an analytical model. Our model captures periods of connectivity and disconnection with a two-state loss process. To determine the effectiveness of these schemes, we evaluated the probability of successful delivery and the expected number of link-layer transmissions under a limited number of transmission attempts.

We introduced the spaced hop-by-hop scheme and observed that it dominates the end-to-end scheme for all values of disruption duration. Moreover, the spaced scheme dominates the basic hop-by-hop scheme for all values of disruption duration where the correlation of the loss process is positive (i.e., if the disruption durations are relatively long).

We are currently extending our analysis; preliminary results suggest that there is full dominance of spaced hop-by-hop over end-to-end transport in terms of probability of delivery, expected number of link transmissions, and end-to-end latency. The details of these results are to appear in a forthcoming paper.

7. ACKNOWLEDGMENTS

The authors would like to thank Pavan Nuggehalli for the inspiring discussions and the valuable feedback.

8. REFERENCES

[1] D. Aguayo, J. Bicket, S. Biswas, G. Judd, and R. Morris. Link-level measurements from an 802.11b mesh network. In *ACM SIGCOMM '04*, New York, NY, USA, Aug. 2004.

[2] A. DeSimone, M. C. Chuah, and O.-C. Yue. Throughput performance of transport-layer protocols over wireless LANs. In *IEEE GLOBECOM '93*, Houston, TX, USA, Nov. 1993.

[3] H. Elaarag. Improving TCP Performance over Mobile Networks. *ACM Computer Surveys*, 34(3), 2002.

[4] S. Floyd, T. Henderson, and A. Gurtov. The NewReno Modification to TCP's Fast Recovery Algorithm. RFC 3782 (Proposed Standard), Apr. 2004.

[5] I.I. Gichman and A.V. Skorokhod. The Theory of Stochastic Processes II. Springer, May 2004

[6] I. Gitman. Comparison of Hop-by-hop and End-to-end Acknowledgment Schemes in Computer Communication Networks. *IEEE Trans. on Comm.*, 24(11), 1976.

[7] S. Heimlicher, P. Nuggehalli, and M. May. End-to-end vs. Hop-by-hop Transport. In *ACM SIGMETRICS '07 Student Workshop*, San Diego, CA, June 2007. To app.

[8] S. Heimlicher, R. Baumann, M. May, and B. Plattner. The Transport Layer Revisited. In *IEEE COMSWARE '07*, Bangalore, India, Jan. 2007.

[9] IEEE. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications. In *IEEE Std 802.11-1999*, 1999.

[10] T. Issariyakul and E. Hossain. Analysis of end-to-end performance in a multi-hop wireless network for different hop-level ARQ policies. In *IEEE GLOBECOM '04*, volume 5, Nov. 2004.

[11] V. Jacobson and M. Karels. Congestion Avoidance and Control. In *ACM SIGCOMM '88*, Aug. 1988.

[12] D. B. Johnson and D. A. Maltz. Dynamic source routing in ad hoc wireless networks. In *Mobile Computing*, volume 353. Kluwer, 1996.

[13] P. Juang, H. Oki, Y. Wang, M. Martonosi, L.-S. Peh, and D. Rubenstein. Energy-Efficient Computing for Wildlife Tracking: Design Tradeoffs and Early Experiences with ZebraNet. In *ASPLOS-X*, San Jose, CA, USA, Oct. 2002.

[14] J. C. Kuo and W. Liao. Hop count distribution of multihop paths in wireless networks with arbitrary node density: Modeling and its applications. *IEEE Trans. on Vehic. Techn.*, 56, July 2007.

[15] V. Lenders, J. Wagner, and M. May. Analyzing the Impact of Mobility in Ad Hoc Networks. In *ACM REALMAN*, Florence, Italy, May 2006.

[16] P. P. Mishra, H. Kanakia, and S. K. Tripathi. On hop-by-hop rate-based congestion control. *IEEE/ACM Trans. on Netw.*, 4(2), Apr. 1996.

[17] C. Perkins, E. Belding-Royer, and S. Das. Ad hoc On-Demand Distance Vector (AODV) Routing. RFC 3561 (Experimental), July 2003.

[18] C. E. Perkins and P. Bhagwat. Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers. In *ACM SIGCOMM '94*, New York, NY, USA, 1994.

[19] C. E. Perkins and E. M. Royer. Ad-hoc on-demand distance vector routing. In *IEEE WMCSA '99. IEEE*, Washington, DC, USA, 1999.

[20] J. Postel. Internet Protocol. RFC 791 (Standard), Sept. 1981.

[21] J. Postel. Transmission Control Protocol. RFC 793 (Standard), Sept. 1981.