Inter-Regional Messenger Scheduling in Delay Tolerant Mobile Networks

Khaled A. Harras and Kevin C. Almeroth
Department of Computer Science
University of California, Santa Barbara
{kharras, almeroth}@cs.ucsb.edu

Abstract

The evolution of wireless devices along with the increase in user mobility have created new challenges such as network partitioning and intermittent connectivity. These new challenges have become apparent in many situations where the transmission of critical data is of high priority. Disaster rescue groups, for example, are equipped with numerous devices which constantly gather and transmit various forms of data. The challenge of establishing communication between groups of this type has led to an evolutionary form of networks which we consider in this paper, namely, Delay Tolerant Mobile Networks (DTMNs). Nodes in DTMNs usually form clusters that we define as regions. Nodes within each region have end-to-end paths between them. Both regions, as well as nodes within a region, can be either stationary or mobile. For such environments, we propose using a dedicated set of messengers that relay message bundles between these regions. Our goal is to understand how messenger scheduling can be used to improve network performance and connectedness. We develop several classes of messenger scheduling algorithms which can be used to achieve inter-regional communication in such environments. We use simulation to better understand the performance and tradeoffs between these algorithms.

1 Introduction

The evolution of wireless devices, such as laptops, personal digital assistants (PDAs), and cell phones, is a fact we are all currently witnessing. These devices are increasingly being used to accomplish basic day-to-day tasks. The growing dependence on these devices, along with the high mobility of users, has increased the need to be connected in all places at all times. This increase in user demand has spurred the development of numerous applications that run in new network environments. Examples of these environments include: satellite networks, planetary and interplanetary communication, military/tactical networks, disaster response, and other forms of large-scale mobile networks. Such environments have created a number of new challenges for network designers to solve. These new challenges include, but are not limited to, network partitioning, intermittent connectivity, large delays, high deployment cost, and the absence of an end-to-end path.

These new challenges have spurred much research in mobile environments. Most of the research in this area focuses on that of Mobile Ad Hoc NETworks (MANETs), which are multi-hop networks in which nodes act as routers and cooperate to maintain end-to-end connectivity. Most of the work in MANETs, therefore, is targeted at solving the routing problem [9], [10], [13], [6], [15], [14]. MANETs, however, fail to address all of the emerging challenges listed earlier since they focus on scenarios where an end-to-end path *must exist* from a source to a destination.

New areas such as Delay Tolerant Networks (DTNs) and partially-connected mobile networks, have consequently emerged to address these new challenges. The DTN community [1], has proposed a general architecture [5], that addresses routing issues [8] for networks in extreme environments. On the other hand, partially-connected mobile networks are targeted towards solving these problems in more specific scenarios and environments. Some examples include DataMULEs [16], Message Ferrying [19], [18] and Epidemic Routing [17]. The general DTN architecture works well by trying to address common problems among different networks in challenged environments. Conversely, other specialized solutions are more tailored to the specific problem they try to solve, and therefore, are more customized to their designated environments. Other problems and issues that were previously overlooked, however, require further research.

To demonstrate the problems we are concerned with in this paper, we consider several scenarios such as disaster relief efforts and field hospitals [2], battlefields [11], and remote disconnected villages [3], [4]. In these scenarios, we observe a class of Delay Tolerant Mobile Networks (DTMNs) [7] where nodes form clusters such that a communication path exists between any two nodes *within* each cluster. Nodes in *different* clusters, however, cannot communicate except through long-range and high-power wireless or satellite networks. This limitation often occurs because existing infrastructure is either destroyed (e.g. after a

hurricane or earthquake), or simply do not exist (e.g. battlefield areas or remote disconnected villages). Also, if the cost of providing these forms of communication is too high or the type of data cannot be transmitted over such networks, these clusters would then need to use other communication paths and methods. This issue becomes particularly important in cases where large amounts of multimedia data, such as images of disaster areas or video surveillance clips, must be transmitted between clusters.

For such scenarios, we envision a new network environment comprised of *regions* and *messengers*. A region is defined as a cluster of nodes having an end-to-end path between any two nodes in the cluster. Since regions are assumed to be disconnected from each other, we propose using a dedicated set of messengers that relay messages between regions. Each region generates large amounts of data that can be grouped into *bundles* [5], which are then relayed to other destination regions. Multiple messengers would provide fault tolerance and faster delivery of message bundles. The regions could either be mobile, as in search-and-rescue groups or military battalions, or stationary, as in field hospitals or remote disconnected villages. In these environments, we shift the focus from routing [8] and path discovery [18], to *messenger scheduling*. We study how messenger scheduling can be used to improve network performance and connectedness. Furthermore, there are a number of scheduling algorithms (e.g. batching) that can be drawn from marginally related areas, and used to provide insight into how best to schedule messengers.

The contribution of this paper is to develop classes of messenger scheduling algorithms that can be used to achieve inter-regional communication in the environments described above. The novelty of our work is in two main areas. First, we consider environments where regions are mobile and dynamic in nature, as opposed to single stationary or mobile nodes. Second, we introduce the idea of using dedicated messengers under different scheduling strategies rather than discovering and maintaining routing paths in these kinds of networks. We propose different messenger ownership and scheduling schemes which result in a two-dimensional set of solutions representing various scheduling algorithms for messengers. Using simulation, we evaluate these algorithms under different network conditions and study their performance in terms of delay, cost, and efficiency. Our main goals are to gain a better understanding of the challenges involved in the complex network environments we introduce, and to identify which scheduling algorithms are most suitable for different network environments.

The remainder of this paper is organized as follows. Section II introduces related work. Section III presents our system architecture and scheduling algorithms. The simulation environment and results are described in Section IV. Finally, we present our conclusions in Section V.

2 Related Work

Solutions have been proposed to address the various scenarios where an end-to-end connection cannot always be assumed. These solutions can be divided into two main categories: partially-connected mobile networks and Delay Tolerant Networks (DTNs). Most, if not all, solutions in partially-connected mobile networks rely on some form of store and forward approach, and they can be generally classified into *random* and *non-random* schemes. In a random scheme, nodes passively wait until they come within range of another node as a result of their normal mobility pattern. A non-random scheme assumes the ability of a node to actively change its mobility pattern or trajectory, based on some information about various nodes in the network.

Examples of random schemes in partially-connected mobile networks are Epidemic Routing [17] and DataMules [16]. In Epidemic Routing, Vahdat and Becker [17] introduce a flooding-based routing protocol for partially-connected mobile networks. Nodes in the network continuously exchange copies of the messages that they do not have, until the messages reach their intended destination(s). Shah et al. [16] introduce DataMULEs, where low powered static sensors are sparsely deployed to gather various forms of data, and then a mobile entity, a "mule", randomly travels among these sensors to collect the data that they gathered. We observe that these random schemes are too expensive, impractical, or insecure for the scenarios mentioned in Section I. They also do not consider using or scheduling dedicated messengers between regions as we do in this paper.

Examples of non-random schemes in partially-connected mobile networks include the work by Li and Rus [12] as well as Message Ferrying [18]. Li and Rus propose a scheme where nodes actively change or modify their trajectory to help create a path which enables message delivery to the destination more quickly [12]. While this adaptive node trajectory feature is useful in situations where all the nodes can be controlled and belong to one entity, extending their work to support multiple messages is difficult. Zhao and Ammar introduce the idea of Message Ferrying and study its performance with stationary and mobile nodes in the network [19]. They also explore controlling the mobility of multiple ferries by studying algorithms that compute different routes for the ferries [18].

We note at this point that all these non-random schemes are limited when compared to our work. Either they do not support multiple messages simultaneously [12], or consider only a single ferry [19], or offer only multiple predetermined fixed routes for ferries with stationary nodes [18]. In our paper, we take advantage of the possibility of having clusters of nodes that could communicate among themselves, and examine cases where both the nodes and regions are mobile. We avoid maintaining routing information that would be outdated and difficult to maintain in such com-

plex situations. We also focus on studying different point-to-point scheduling algorithms.

In the area of Delay Tolerant Networks (DTNs), Fall provides a generalized overlay architecture as an attempt to achieve inter-operability between heterogeneous networks deployed in extreme environments [5]. These networks usually lack continuous connectivity and suffer from potentially long delays. Jain et al. expand the DTN work by studying routing issues in extreme environments [8]. In our previous work, we examine a special case of DTNs, which we introduced as Delay Tolerant Mobile Networks (DTMNs) [7]. These networks are comprised of large-scale sparse mobile networks where no end-to-end path is assumed to exist between any two nodes in the network. We study different controlled flooding schemes to minimize the consumption of network resources and compare their performance in such environments.

The drawback of the DTN architecture [5] is that it only provides general guidelines and a framework to shed light on common problems that exist in different challenged networks and extreme environments. Additionally, it mainly focuses on routing issues in extreme environments under the assumption of different knowledge "oracles" [8]. Similarly, in DTMNs [7], we do not take advantage of clusters of nodes and do not use *dedicated messengers* for communication. Our work differs from previous DTN-related research in that DTN gateways in our system are the mobile messengers which we propose in the following section. These dedicated messengers require efficient scheduling algorithms to make the best use of them. As a result, we address the issue of messenger scheduling, without having predefined fixed paths, in order to adapt to the varying demand for messengers in DTMN environments.

3 System Architecture

Our goal in this section is to set the boundaries for our work and to define the concepts and ideas we introduce. We first briefly describe Delay Tolerant Mobile Networks (DTMNs). We then define and clarify the notion of *regions* and *messengers*. Finally, we discuss our different messenger ownership schemes and scheduling strategies and show how these two sets can be merged to create the scheduling algorithms we study.

3.1 Delay Tolerant Mobile Networks

The work presented in this paper uses DTMNs [7] as the underlying network environment. We study our scheduling algorithms in this environment and introduce the concepts of regions and messengers. In this section, we give a brief overview of the basic DTMN architecture to help introduce the notion of regions, messengers, and our scheduling algorithms.

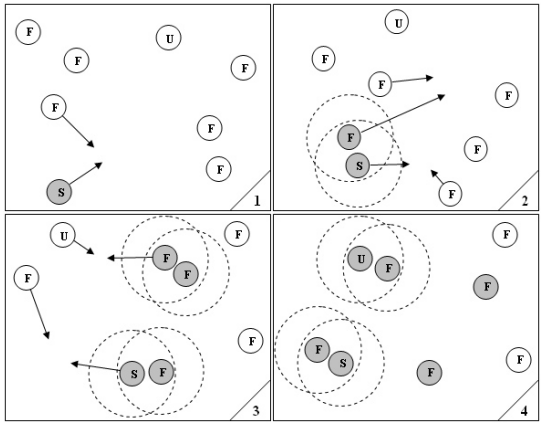


Figure 1. An example of message delivery in a DTMN. "S" is a source node, "F" are forwarder nodes, and "U" is the ultimate destination node. Shaded nodes represent those that have received copies of the message.

Our concept of DTMNs focuses on a special kind of Delay Tolerant Networks (DTNs) where all nodes are assumed to be mobile, and where no end-to-end path necessarily exists between any two nodes in the network. Each node is viewed as an independent region with respect to the DTN architecture [5]. Each node in this case acts as a DTN gateway to perform overlay bundle relaying of messages. There are two basic assumptions regarding nodes in DTMNs. First, nodes are "blind", i.e. the nodes in the network do not have any information regarding the state, location, or mobility patterns of other nodes. Second, nodes are "autonomous", i.e. each node has independent control over itself and its movement.

Under the assumptions of blindness and autonomy, the most common technique for message relaying and delivery is to use flooding, e.g. Epidemic Routing [17]. Figure 1 shows the basic way in which a given message propagates through the network from the source node, *S*, to an ultimate destination node, *U*. This is achieved through the aid of other forwarder nodes, *F*, that relay the message until it reaches the intended destination. In previous work, we focused on controlling these floods to minimize the consumption of network resources, and examined different schemes and combinations of these schemes [7]. In this paper, however, we look at other forms of DTMNs where the assumptions of blindness and autonomy no longer hold. The components and architecture of this new class of DTMNs are introduced in the following section.

3.2 Regions and Messengers

In DTMNs, there are many cases where clusters of nodes or hot spots are formed. These clusters are either mobile or stationary. More importantly, these clusters have end-to-end paths between nodes within each cluster. These clus-

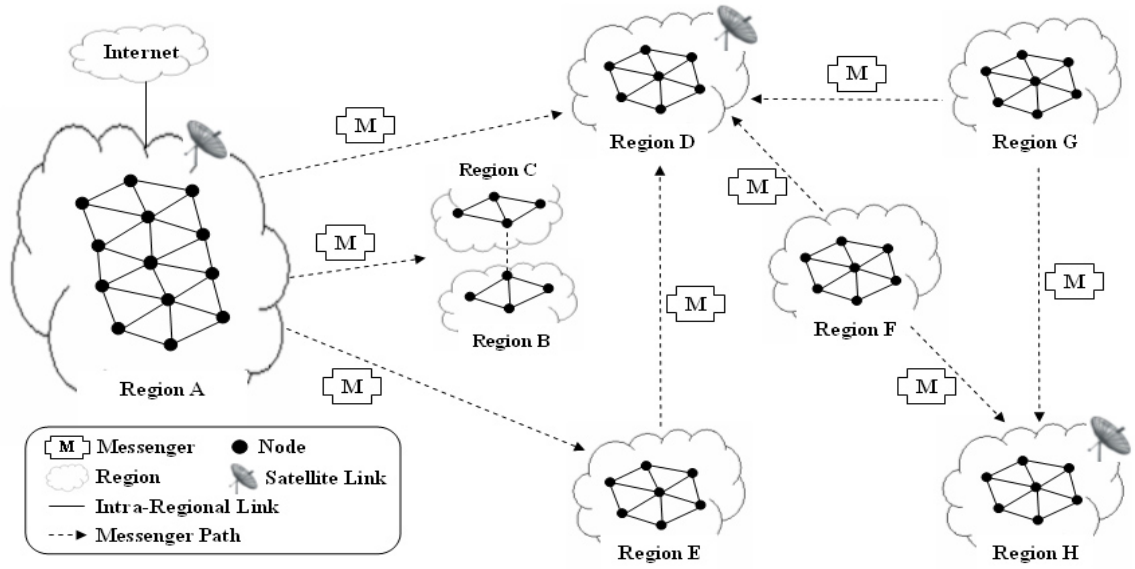


Figure 2. Diagram demonstrating regions and messengers in a Delay Tolerant Mobile Network (DTMN).

ters, however, require some form of communication between them. To satisfy this requirement, this section discusses the notion of *regions* and *messengers* in DTMNs.

Figure 2 demonstrates the concepts of *regions* and *messengers*. A region is defined as an intra-connected cluster of nodes or hot-spots that exist in extreme environments. A region could represent a military battalion, a remote village, or a disaster response team. Region formation, and the membership of a node in a region, are administrative decisions that are taken depending on the nature of the underlying system. The transport layer protocols used for intra-regional communication is a decision left to the region itself; each region selects what best suits its class of applications and requirements. Since regions are defined to be mobile, they can dynamically divide or merge. For example, in Figure 2, Regions B and C are about to merge into a single region. For simplicity, however, we focus in this paper only on cases where regions are mobile, but do not divide or merge.

Messengers, shown in Figure 2, are the entities responsible for achieving inter-regional communication. They are dedicated to carrying *bundles* of data between regions. A data bundle is a collection of messages that needs to be sent from one or more nodes in a given region to nodes in another region. These messengers could be helicopters, drones, robots, motorbikes, busses, trains, or even planes depending on the application and environment of operation.

A key requirement of our scheduling algorithms is that messengers know where they are going, that the location of other regions and their movement are known. To begin with, each region can receive information regarding its location through Global Positioning System (GPS). It is certainly reasonable to assume that at least one node in each region will have GPS, which is sufficient for a region to know its own location. The current location of a region could then be transmitted to the messengers through various long-range communication technologies. This requirement is not unreasonable for any of the scenarios or environments we consider. Military battalions or disaster recovery teams, for example, will have at least some basic communication facility for emergency communication. Our contention however, is that this facility is insufficient for large file transmission or secure communication, but is certainly sufficient for transmitting region location coordinates.

At this point, we note that we are not addressing routing issues between regions in extreme environments. Routing has been studied in DTN environments [8]. Path determination for ferries in DTNs has also been examined [18]. Routing in this environment could be solved using a number of existing algorithms. For example, the image in Figure 2 could be reduced to a graph with each region represented by a single node and then finding the shortest paths among this set of nodes. Edge weights would determine the cost of a given path between two regions. The cost, could be a function of the distance, difficulty of traversing the terrain, or cost of operating the messenger. These issues are dealt with in other work and are outside the scope of this paper. We focus instead on the question of how to schedule the messengers that will travel between regions.

We choose to look at such DTMN environments in a distributed fashion, where messengers are scheduled independently in each region. We then propose and study different assignment and scheduling strategies for messengers. Merging these two strategies leads to scheduling algorithms that operate dynamically and can be extended to adapt to varying network resource demands.

3.3 Messenger Ownership and Scheduling

In this section, we present the different messenger ownership schemes and scheduling strategies we propose for achieving inter-regional communication in delay tolerant mobile networks. We propose six different strategies by which messengers are assigned and scheduled in a DTMN environment. These approaches are summarized by the grid shown in Figure 3. The Y-axis in the grid represents the ownership of the messenger, while the X-axis describes the time basis upon which a messenger will move to transmit a message to a given destination. The grid can be expanded to include other possibilities resulting in a larger spectrum of solutions. We find, however, that the approaches presented in Figure 3 represent the major points within this spectrum.

3.3.1 Messenger Ownership

The Y-axis in Figure 3 describes how messengers in a DTMN system are assigned to different regions in the network. We present two alternative assignment strategies: regional messengers and independent messengers.

Regional Messengers mean that each messenger is *owned* by a certain region in the system. The messenger belongs to a given region which is the source region of a given bundle of messages. This messenger can carry these bundles from the source region to a destination region. The regional messenger can also carry messages from a destination region back to the source (owner) region. Once a messenger reaches a destination region, it delivers the message bundle and then immediately returns to its owner region. It can only carry messages from this destination region that are addressed to its source (owner) region. Regional messengers can, therefore, only carry messages that are either sent by its owner, or destined to its owner region from the destination region to which it was initially sent.

Independent Messengers mean that each messenger in the system is not owned by any region. The messenger is managed by the region where it currently resides. This concept can be viewed as temporary ownership by a given source region that ends once the source region sends a bundle to a new destination. The ownership then transfers to the destination region. Once a messenger reaches a destination region, it delivers the message bundle and then resides in this new region. Independent messengers, therefore, can basically carry messages from one region to any other region in the network. Regions, on the other hand, do not own any messengers, but simply use the available set of messengers at a given point in time to transmit the bundles it has.

3.3.2 Messenger Scheduling Time

The X-axis in Figure 3 describes the different scheduling strategies for sending messengers. We present three alternative strategies: periodic, on-demand, and storage-based scheduling.

Scheduling Time / Ownership	Periodic	On-Demand	Storage-Based
Regional	RP – Regional Periodic	ROD – Regional On-Demand	RSB – Regional Storage-Based
Independent	IP – Independent Periodic	IOD – Independent On-Demand	ISB – Independent Storage-based

Figure 3. Assignment and Scheduling Strategies.

Periodic time scheduling is equivalent to a shuttle system, where shuttles leave either at a given time or after a fixed period of time. In other words, messengers in a given region are set for departure to a certain destination region at pre-determined times. The messengers are sent whether they have message bundles to transmit or not.

On-Demand time scheduling means that messengers within a given region are sent to another region as soon as the source region has any message to send. In the periodic or storage-based approaches, messages from multiple sources within the same region are bundled together to be delivered to a given destination. In this case, however, the messenger, assuming one is available, travels as soon as any node within the source region has a message that is required to be sent. This property can be viewed as a high-priority system.

Storage-Based time scheduling falls somewhere between periodic and on-demand scheduling. A messenger in this case starts to move towards its destination as soon as it has a predefined storage capacity filled by messages that need to be delivered. Tweaking the storage limit in this case affects the overall performance and efficiency of messenger usage in such networks.

4 Evaluation

The primary goal for our evaluation is to observe the tradeoffs in the performance of our scheduling strategies. We also identify which scheduling and ownership schemes are more appropriate for different network conditions and environments. We first describe our simulation setup and environment. We then give an analysis that clarifies and sets bounds for some of the metrics we consider. Finally, we discuss and analyze our simulation results.

4.1 Simulation Environment

In our simulation setup, we have n regions randomly distributed over a $10km$ x $10km$ terrain. This area is realistic for a rescue team or military battalion, for example. We also have k messengers that are equally distributed over the regions in the system. The regions are mobile, and use a *modified* random way-point mobility model. We particularly avoid the major problem of the network slowing down in the conventional random waypoint model. Also, while random way-point may not be the best model for individual

Table 1. Simulation Parameters

Parameter	Value Range	Nominal Value
Terrain	$10km$ X $10km$	$10km$ X $10km$
Simulation Time	1 hour to 24 hours	12 hours
# of Regions (R)	3 to 10	5
# of Messengers	R to R^3	R^2
Message Rate	30 to 150 messages/h	100
Message Pattern	Uniform, Exponential	N/A
Region Speed	$4km/h$ to $40km/h$	N/A
Messenger Speed	$50km/h$ to $90km/h$	N/A
Traffic Pattern	Distributed Uniform or Many-to-one	Distributed Uniform
Scheduling Alg.	Periodic, SB or OD	N/A
Ownership	Regional, Independent	N/A
Periodic Time	5 mins to 120 mins	30 mins
Storage Limit	10 to 1000 messages	50 messages

mobile nodes, we believe that it suits the regional mobility patterns with which we are concerned (e.g. in disaster relief). Each region has a diameter, r, that determines the size, and transmission range of the region. Finally, we assume that at least one node in each region has GPS, and therefore, messengers ,also equipped with GPS, could dynamically track the location of all regions.

The parameters we believe have the most impact on our system are summarized in Table 1. Generally speaking, the speed of messengers is larger than that of regions. We observed that the actual speed range for regions and messengers does not have much impact on the *relative* performance of the scheduling strategies. This is because we are mainly concerned with the tradeoffs between these strategies rather than their absolute performance. We study the impact of different ownership and scheduling strategies on our metrics. We also consider uniform and exponential message generation patterns from each region in our system.

We perform experiments over different inter-regional traffic patterns. We particularly study *many-to-one* traffic and *distributed-uniform* traffic patterns. In the many-to-one pattern, many regions send messages to one region in the system, while in the distributed-uniform pattern, all regions have an equal probability of sending messages to any other region. In this way, we satisfy scenarios where several regions need to send data or updates to a central region (e.g. a battlefield command center) or all regions simply distribute their data to all other regions (e.g. search and rescue teams). We also vary the message generation rate in our system to observe the performance under different loads. Other parameters that impact our system are shown in Table 1.

We consider three metrics in evaluating the scheduling strategy tradeoffs in our system. These metrics are:

- **Delay:** measured by the average amount of time taken by a message to go from its source region to a destination region.
- **Cost:** measured by the the total number of trips the messengers have taken after a given amount of time.
- **Efficiency:** measured by the average number of messages carried and delivered in each trip. This metric measures messenger usage efficiency.

4.2 Simulation Results

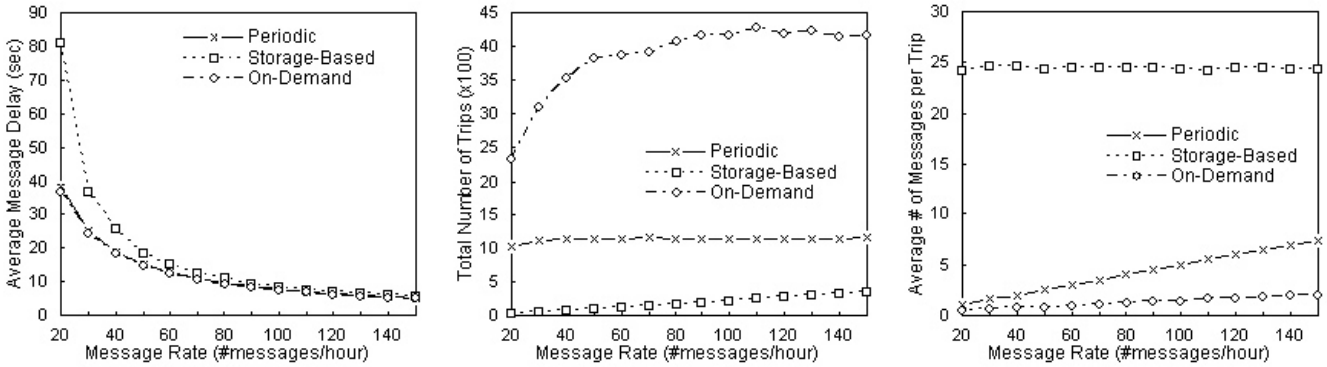
In our evaluation, we conducted an extensive set of simulations, but will focus in this section on the subset that best illustrates the key results that lead to our conclusions. We show the tradeoffs between the performance of the different scheduling strategies and ownership schemes that we propose. We first demonstrate the impact of both varying the message generation rate as well as changing the number of regions on our scheduling strategies, under the two different ownership schemes. Next, we show the operation of our scheduling schemes under different network traffic patterns. Finally, we briefly discuss other result sets that we obtained which provide more insight into our system. All measurements are taken with respect to the three metrics of delay, cost, and efficiency. Nominal values of the parameters in Table 1 are used for all experiments, except for those parameters being tested. Each point in our results is taken as an average of 20 different simulation seeds.

4.2.1 Impact of the Message Rate

In this section, we demonstrate the performance of the periodic, storage-based, and on-demand scheduling strategies while varying the message generation rate. Figure 4 shows the impact of changing the message generation rate on our scheduling strategies using the regional ownership scheme. Figure 5 shows the same results for the independent ownership scheme.

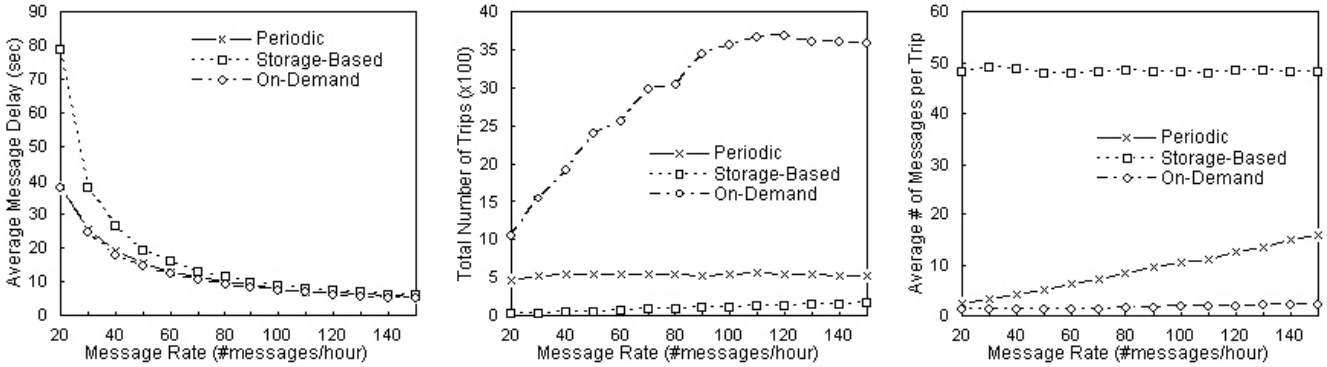
In general, Figures 4 and 5 show similar relative behavior in terms of the scheduling strategy performance. We observe that delay in both Figures 4(a) and 5(a) decreases exponentially as the message rate increases. The reason for this result is that at high message rates, different components of the total delay cancel each other: queuing time and messenger wait time. At low message rates, the storage-based scheme suffers from a much larger delay when compared to the other two schemes since it takes longer to fill the storage limit with lower rates.

With respect to cost, shown in Figures 4(b) and 5(b), we see an increase in the cost incurred by the storage-based and on-demand schemes. The figure also shows that the results for the periodic scheme remain constant. This result occurs because the periodic scheme operates independently of the message generation rate; messengers move after the periodic time expires regardless of the number of messages it carries, and therefore, execute the same number of trips. The on-demand and storage-based schemes are



(a) Delay (b) Cost (c) Efficiency

Figure 4. Impact of the message generation rate on the scheduling strategies under the *regional* ownership scheme.



(a) Delay (b) Cost (c) Efficiency

Figure 5. Impact the message generation rate on the scheduling strategies under the *independent* ownership scheme.

sensitive to the change in message rate, since a higher message rate would cause more on-demand trips, or will reach the storage limit more quickly.

The efficiency of the system, shown in Figures 4(c) and 5(c), increases for the periodic and on-demand schemes while remaining constant for the storage-based scheme. This result occurs because the efficiency of the storage-based scheme is fixed since the messenger does not move unless its storage limit is reached. However, for the other two schemes, as the message rate increases, more messages can then be loaded when a messenger is waiting (periodic), or more messages can be carried back to the owner region (on-demand).

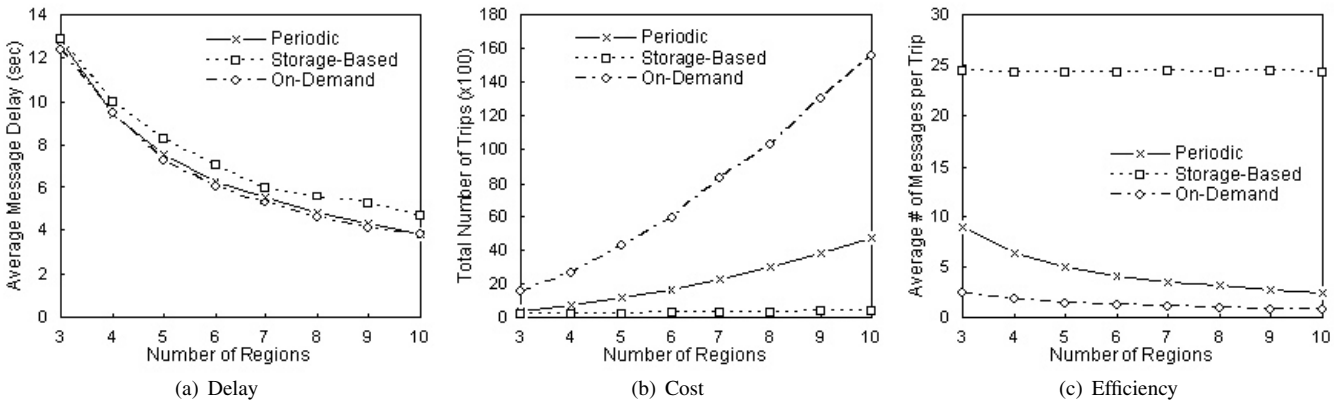
We observe that under both ownership schemes, the storage-based strategy seems to have the least cost and highest efficiency, while the on-demand strategy has the highest cost and the lowest efficiency. The on-demand strategy, however, incurs less delay for low message rates. Even though the general behavior of the strategies looks the same under different ownership schemes, the difference in terms of cost and efficiency is large in both cases. This observation can be clearly seen when we compare the Y-axis scales for those two metrics under different ownership schemes. We find that the regional scheme (Figure 4) incurs approximately double the cost and half the efficiency when compared to the independent scheme (Figure 5), for all scheduling strategies. This result is due to the extra trip each messenger must take back to its owner region after delivering messages.

4.2.2 Impact of the Number of Regions

We now compare the different scheduling strategies in our system while keeping the message rate constant and changing the number of regions, and consequently, the total number of messengers[1]. Figure 6 shows the result of changing the number of regions on our scheduling strategies using the regional ownership scheme. Figure 7 shows the same results for the independent ownership scheme.

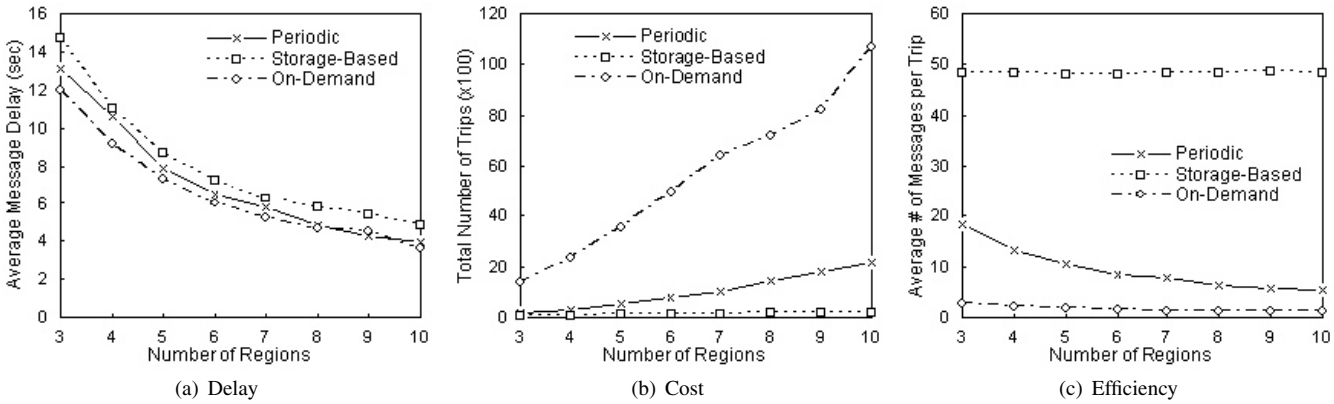
In general, Figures 6 and 7 show similar relative behavior in terms of scheduling strategies performance according to our metrics under both regional and independent ownership schemes. However, when focusing on the axis scales, we see that they have different absolute behavior. We ob-

[1]See Table 1 for the relationship between the number of messengers and the number of regions.



(a) Delay (b) Cost (c) Efficiency

Figure 6. Impact of the number of regions on the scheduling strategies under the *regional* ownership scheme.



(a) Delay (b) Cost (c) Efficiency

Figure 7. Impact of the number of regions on the scheduling strategies under the *independent* ownership scheme.

serve that delay in both Figures 6(a) and 7(a) decreases as the message rate increases. The reason for this result is that as the number of regions increases, the number of messengers in each region also increases. This fact, along with a fixed message rate, means that there are more messengers available more often to be used by the messages queued in each region.

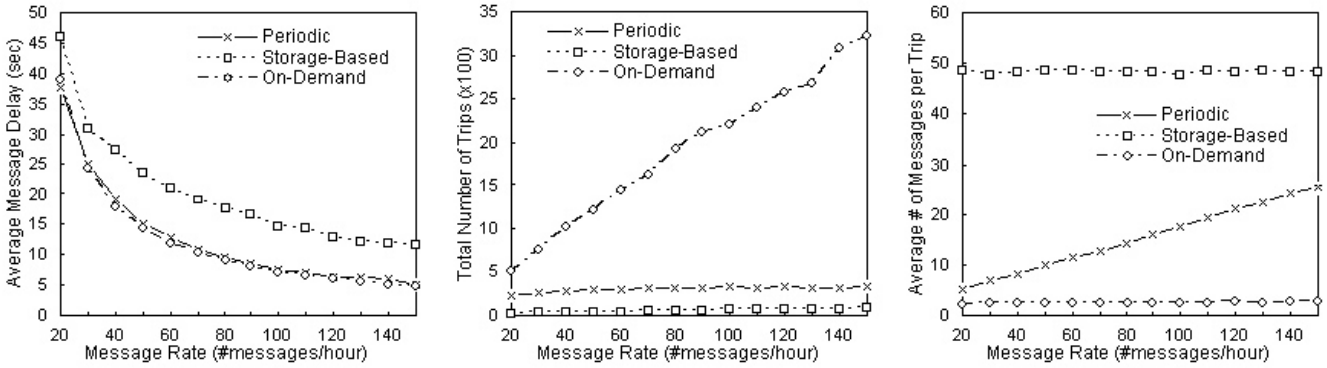
With respect to cost, shown in Figures 6(b) and 7(b), we see an increase in the cost for all scheduling strategies. This result occurs because, since we have more messengers to compensate for the increase in the number of regions, a resulting increase in the total number of trips is expected. Finally, the efficiency of the system, shown in Figures 6(c) and 7(c), decreases for the periodic and on-demand scheduling strategies while remaining constant for the storage-based strategy. The reasoning is the same as that described in the previous section. If a messenger moves only when its storage limit is reached, it maintains a constant efficiency level. The efficiency for the on-demand and periodic schemes decreases, however, because the same number of messages are divided over an increasing number of destinations. The overall result is a decrease in the number of messages each messenger carries when traveling to a given destination.

Similar to the behavior in the previous section, and for the same reasons, the storage-based strategy has the least cost and the highest efficiency, while the on-demand scheme has the highest cost and the lowest efficiency. Again, we observe that the regional scheme (Figure 6) incurs approximately double the cost and half the efficiency when compared to the independent scheme (Figure 7), for all scheduling strategies.

4.2.3 Impact of the Traffic Pattern

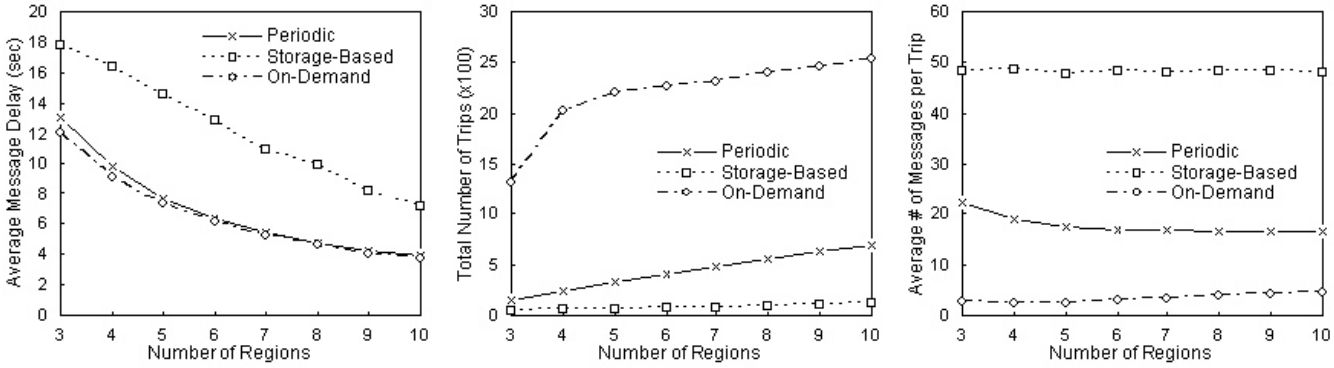
So far, all results that we have shown have been based on simulations using a distributed uniform traffic pattern. In this section, we demonstrate the impact of adopting a many-to-one pattern and show a subset of our results.

Figure 8 shows the impact of changing the message generation rate on our scheduling strategies using the independent ownership scheme. Figure 9 shows the impact of changing the number of regions on our scheduling strategies also under the independent ownership scheme. While the general performance looks familiar, the results need to be compared to those in Figures 5 and 7, respectively, to



(a) Delay (b) Cost (c) Efficiency

Figure 8. Impact of the *message rate* on the scheduling strategies with independent ownership and a many-to-one traffic pattern.



(a) Delay (b) Cost (c) Efficiency

Figure 9. Impact of the *No. of regions* on the scheduling strategies with independent ownership and a many-to-one traffic pattern.

grasp the impact of using a many-to-one traffic pattern as opposed to a distributed uniform one. In general, there is a reduction in the cost incurred and an increase (except for the storage-based scheme) in efficiency in Figures 8 and 9 when compared to Figures 5 and 7, respectively.

The intuition behind this variation in results is as follows. In general, under a many-to-one scheme operating with independent ownership, the environment beings to look like a basic command center or sink where most of the messengers gradually accumulate. When these messengers are sent to their destinations, they gather as many messages as possible (except for storage-based), from those queued at these destination regions, that need to be sent back to the sink region. This behavior works well under average to high message generation rates. However, if the message generation rate is very low, or the sink does not produce any messages at all, the independent scheme in this case could cause starvation when all the messengers accumulate at the sink and are rarely, if ever, sent to the other regions. The solution in such cases is simple: a periodic timer could be used to re-distribute messengers over the regions in the system.

We do not show results for the regional ownership scheme because the overall performance tends to looks similar to that of the distributed uniform traffic pattern. The reason for this result is that each region owns a number of messengers that must return to it upon message delivery.

4.2.4 Other Result Sets

So far, we have presented a subset of results for cases we believe are common in the applications we consider. However, we conducted numerous other simulations that we do not show in this paper due to the rareness with which these situations are likely to occur. Extreme conditions with large message bursts or very low message generation rates are good examples of rare cases. The results in such conditions generally show that on-demand tends to behave reasonably well with very low rates since messengers are sent only when a message is generated, which rarely occurs. Under large message bursts, the periodic or storage-based schemes seem to operate well.

Varying the number of messengers while keeping the number of regions fixed is an example of another result we evaluated. In this paper, we only show results where the number of messengers increases with the increase in regions because we believe that the number of messages de-

ployed will be proportional to the expected number of regions. However, if the number of regions were to increase while maintaining the same number of messengers, delay would increase due to the increase in the resulting queuing time of messages at the regions. This result occurs because there are more destinations that need to be served with the same number of messengers.

Finally, we note that prior to the subset of results we presented, we studied the impact of changing the periodic time and storage limit values on both the regional and independent ownership schemes. This is because the performance of the periodic and storage based schemes relies heavily on the choice of these values. Graphs for these results are not shown since they do not contribute much towards our evaluation goals.

5 Conclusions

In this paper we have studied the idea of using a dedicated set of messengers for message delivery in Delay Tolerant Mobile Networks (DTMNs). We have focused on DTMNs where nodes form clusters that we have defined as regions, which are disconnected from each other. Messengers are then used to communicate between these regions.

We have introduced two messenger ownership schemes, regional and independent, as well as three scheduling strategies for message delivery in our system, periodic, storage-based, and on-demand. We have studied the tradeoffs between each of those schemes under different environments and network conditions. Our results have demonstrated that the choice of a particular scheme ultimately depends on the environment under which it is deployed. While the on-demand promises the least delay, it comes with a large cost and low efficiency. Periodic, on the other hand, seems to be a reasonable solution as it provides a close to on-demand delay with reasonable cost and efficiency results. Selecting an appropriate periodic time, though, is a challenge. Finally, storage-based generally gives the highest efficiency and least cost but requires tweaking the storage limit, and comes at the expense of high delay especially at low message generation rates. We have basically shown that our algorithms are efficient when the proper ownership scheme and scheduling strategy are matched with the traffic patterns and operational characteristics of the system.

Our future work includes studying environments where messengers are destroyed or lost, and therefore, require the deployment of end-to-end reliability mechanisms for message delivery. We also plan to investigate more intelligent and adaptive solutions where messengers can dynamically pick a scheduling strategy, or even a mix of strategies, to enable them to adapt to changing network conditions.

References

[1] DTNRG. Delay Tolerant Networking Research Group. http://www.dtnrg.org/.

[2] University of South Florida: Center for robot-assisted search and rescue. http://crasar.csee.usf.edu/.

[3] Technology and Infrastructure for Emerging Regions. http://tier.cs.berkeley.edu/.

[4] First Mile Solutions. http://www.firstmilesolutions.com/.

[5] K. Fall. A Delay-Tolerant Network Architecture for Challenged Internets. In *ACM SIGCOMM*, Karlsruhe, Germany, August 2003.

[6] Z. Haas and M. Pearlman. The Performance of Query Control Schemes for Zone Routing Protocol. In *ACM SIGCOMM*, pages 167–177, Vancouver, Canada, August 1998.

[7] K. Harras, K. Almeroth, and E. Belding-Royer. Delay Tolerant Mobile Networks (DTMNs): Controlled Flooding Schemes in Sparse Mobile Networks. In *IFIP Networking*, Waterloo, Canada, May 2005.

[8] S. Jain, K. Fall, and R. Patra. Routing in a Delay Tolerant Network. In *ACM SIGCOMM*, Portland, OR, August 2004.

[9] D. Johnson and D. Maltz. *Dynamic Source Routing in Ad Hoc Wireless Networks*, volume 353. Kluwer Academic Publishers, 1996.

[10] Y.-B. Ko and N. Vaidya. Location-Aided Routing (LAR) in Mobile Ad hoc Networks. In *ACM MobiCom*, pages 66–75, Dallas, TX, October 1998.

[11] E. Krotkov and J. Blitch. The Defense Advanced Research Projects Agency (DARPA) Tactical Mobile Robotics Program. *The International Journal of Robotics Research*, 18(7):769–776, July 1999.

[12] Q. Li and D. Rus. Sending Messages to Mobile Users in Disconnected Ad-Hoc Wireless Networks. In *ACM MobiCom*, pages 44–55, Boston, MA, August 2000.

[13] C. Perkins. Ad-hoc On-Demand Distance Vector Routing. In *IEEE Workshop on Mobile Computing Systems and Applications*, pages 90–100, New Orleans, LA, February 1999.

[14] C. Perkins and P. Bhagwat. Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers. In *ACM SIGCOMM*, pages 234–244, London, England, October 1994.

[15] E. Royer and C. Toh. A Review of Current Routing Protocols for Ad-hoc Mobile Wireless Networks. *IEEE Personal Communications Magazine*, 6(2):46–55, April 1999.

[16] R. Shah, S. Roy, S. Jain, and W. Brunette. Data MULEs: Modeling a Three-Tier Architecture for Sparse Sensor Networks. In *In IEEE International Workshop on Sensor Network Protocols and Applications*, Anchorage, AK, 2003.

[17] A. Vahdat and D. Becker. Epidemic Routing for Partially Connected Ad Hoc Networks. *Technical Report CS-200006, Duke University*, April 2000.

[18] W. Zhao and M. Ammar and E. Zegura. Controlling the Mobility of Multiple Data Transport Ferries in a Delay-Tolerant Network. In *IEEE INFOCOM*, Miami, FL, March 2005.

[19] W. Zhao, M. Ammar, and E. Zegura. A Message Ferrying Approach for Data Delivery in Sparse Mobile Ad Hoc Networks. In *ACM MobiHoc*, Tokyo, Japan, May 2004.