On the cost/delay tradeoff of wireless delay tolerant geographic routing

Tasiopoulos, Argyrios G ; Tsiaras, Christos ; Toumpis, Stavros

Abstract: In Delay Tolerant Networks (DTNs), there is a fundamental tradeoff between the aggregate transport cost of a packet and the delay in its delivery. We study this tradeoff in the context of geographical routing in wireless DTNs.We first specify the optimal cost/delay tradeoff, i.e., the tradeoff under optimal network operation, using a dynamic network construction termed the Cost/Delay Evolving Graph (C/DEG) and the Optimal Cost/Delay Curve (OC/DC), a function that gives the minimum possible aggregate transportation cost versus the maximum permitted delivery delay.We proceed to evaluate the performance of two known delay tolerant geographic routing rules, i.e., MOVE and AeroRP, a delay tolerant version of the geographic routing rule that selects as next relay the node for which the cost-per-progress ratio is minimized, and finally two novel rules, the Balanced Ratio Rule (BRR) and the Composite Rule (CR). The evaluation is in terms of the aggregate packet transmission cost as a function of the maximum permitted packet delivery delay. Simulations show that CR achieves a cost/delay tradeoff that is overall the closest to the optimal one specified by the OC/DC, while BRR achieves the smallest aggregate transmission costs for large packet delays and a fixed transmission cost model.

DOI: https://doi.org/10.1109/WoWMoM.2012.6263706

Originally published at:
Tasiopoulos, Argyrios G; Tsiaras, Christos; Toumpis, Stavros (2012). On the cost/delay tradeoff of wireless delay tolerant geographic routing. In: WoWMoM 2012, San Francisco, California, USA, 25 June 2012 - 28 June 2012, 1-9.
DOI: https://doi.org/10.1109/WoWMoM.2012.6263706

On the Cost/Delay Tradeoff of Wireless Delay Tolerant Geographic Routing

Argyrios G. Tasiopoulos†, Christos Tsiaras$, and Stavros Toumpis†
† *Department of Informatics, Athens University of Economics and Business, Greece*
$*Department of Informatics, Communication Systems Group, University of Zurich, Switzerland*

***Abstract*—In Delay Tolerant Networks (DTNs), there is a fundamental tradeoff between the aggregate transport cost of a packet and the delay in its delivery. We study this tradeoff in the context of geographical routing in wireless DTNs.**

We first specify the optimal cost/delay tradeoff, i.e., the tradeoff under optimal network operation, using a dynamic network construction termed the Cost/Delay Evolving Graph (C/DEG) and the Optimal Cost/Delay Curve (OC/DC), a function that gives the minimum possible aggregate transportation cost versus the maximum permitted delivery delay.

We proceed to evaluate the performance of two known delay tolerant geographic routing rules, i.e., MOVE and AeroRP, a delay tolerant version of the geographic routing rule that selects as next relay the node for which the cost-per-progress ratio is minimized, and finally two novel rules, the Balanced Ratio Rule (BRR) and the Composite Rule (CR). The evaluation is in terms of the aggregate packet transmission cost as a function of the maximum permitted packet delivery delay. Simulations show that CR achieves a cost/delay tradeoff that is overall the closest to the optimal one specified by the OC/DC, while BRR achieves the smallest aggregate transmission costs for large packet delays and a fixed transmission cost model.

***Keywords*-AeroRP, Cost/Delay Tradeoff, Delay/Disruption Tolerant Networks (DTNs), Dynamic Flows, Evolving Graphs, Geographic Routing, MOVE, Wireless Networks.**

I. INTRODUCTION

Delay Tolerant Networks (DTNs) are networks where the delay in the data delivery is very large, often comparable to the time it takes the network topology to change substantially. Recently, many networks, appearing in disparate settings, have been modeled as DTNs [1], [2], [3].

With respect to designers of traditional networks, designers of DTNs often have an extra degree of freedom, i.e., the delay in the delivery of data. Indeed, this delay can often be traded off in order to improve other metrics. For example, it is shown in [4], [5] that tolerating large delays in wireless networks leads to higher throughputs. In [2] the authors show how delay can be traded off with actual monetary cost, in the context of bulk data transfers over the Internet. The tradeoff existing between, on the one hand, the packet delay and, on the other hand, the consumption of energy and the data storage cost, is often taken into consideration in the context of delay tolerant wireless sensor networks [6], [7].

In this work, we investigate cost/delay tradeoffs involving the aggregate transmission cost and the packet delivery delay in the context of wireless DTNs using geographic routing, i.e., routing where the next hop of a packet is decided taking into account exclusively the location of the destination and the network topology in the immediate neighborhood of the current holder of the packet. Wireless networks that can apply this routing principle are Unmanned Aerial Vehicle (UAV) networks [8], Vehicular Ad Hoc Networks (VANETs) [9], [10], and mobile wireless sensor networks [6], [7], as in these networks nodes can often be equipped with GPS receivers or a related technology.

Related work is discussed in Section II. In Section III we study optimal cost/delay tradeoffs. Section IV focuses on the performance of geographic routing protocols for wireless DTNs, and in particular on the tradeoff between the transmission cost and the packet delivery delay they achieve.

II. RELATED WORK

A. Geographic Routing and Delay Tolerant Networks

Geographic Routing (also called position-based and location based routing) has long been recognized as a key to the creation of scalable routing protocols for use in wireless networks [11], [12], [13], [14]. The basic idea is that the route that a packet takes to go from its source S to its destination D is calculated 'on the fly', as follows: whenever a node A receives the packet, it forwards it to one of its neighbors lying towards the direction of D; the neighbor is selected using a forwarding rule that takes into account only the network topology in the immediate neighborhood of A and the location of D.

Specific geographic routing protocols notably differ on

1) the forwarding rule for selecting the next hop,
2) the precise method used for discovering the location of D, and
3) the mechanism that handles the 'local minimum' situation where the current packet holder appears more suitable to have the packet than all its neighbors, according to the forwarding rule it is applying.

The standard approach for handling 'local minimums' is to immediately initiate a recovery process (such as planarizing the network graph and routing along its arcs, locally modifying the forwarding rule, etc.). More recently, however, geographic routing has been fused with delay tolerant routing [15], [10], [16], [17], [18]. In particular, a packet that arrives at a 'local minimum' is buffered until the topology changes and a suitable next hop becomes available.

For example, in [16] the authors propose MOVE. The basic idea is to forward packets to nodes moving towards the destination (i.e., whose distance to the destination is decreasing with time) and scheduled to pass relatively close to it. In more detail, nodes continuously exchange HELLO message, and a node A will forward a packet destined for node D to its neighbor B if, after a HELLO exchange, any of the following is satisfied:

1) A is either not moving, or moving away from D, while B is moving towards D.
2) A is moving away from D and B is not moving.
3) A, B are moving away from D but B is closer to D.
4) Both A and B are moving towards D, but B is scheduled to pass closer to D than A.

If none of these conditions is fulfilled, then A continues buffering the packet.

In [17] the authors propose AeroRP, a geographic routing protocol for use in DTNs of fast moving airborne vehicles. The basic idea is that a node A holding a packet destined for a destination D should select as the next holder of the packet that node among its neighbors moving towards D that is moving the fastest towards D. In more detail, A defines, for each neighbor B moving towards D, the *Time To Intercept* (TTI) as follows:

$$\text{TTI} = \frac{\Delta d - R}{s_d}, \tag{1}$$

where Δd is the distance of B to D, R is the communication radius of B, and s_d is the relative velocity with which B moves towards D. Node A also calculates its own TTI (if it moves towards D). Node A either forwards the packet to its neighbor with the smallest TTI, or keeps it, if A has the smallest TTI or if both A and all its neighbors are moving away from D.

B. Dynamic Networks

In the context of Network Optimization, our work is on dynamic flows and networks. In contrast to their more common, static counterparts, dynamic networks have properties that change with time, and dynamic flows are functions of time. Ever since their introduction [19], dynamic flows have attracted a steady interest, and an impressive volume of results has accumulated [20], [21], [22]. The typical approach taken is to convert the dynamic problem at hand to a problem involving a static, but typically much larger in size, graph, usually called *space-time graph* or *time-expanded graph*. This is also our approach.

In the context of communication networks, the authors of [23], [24] define *evolving graphs* and use them to solve minimum cost problems and variants in time varying networks; the authors of [25] compute routes for minimizing end-to-end message delivery delay in space-time graphs. In the context of DTNs, the authors of [1] introduce a modification to Dijkstra's algorithm suitable for time varying graphs with propagation delays; in [26] a static graph called *event-driven graph* is introduced and used to devise optimal routing strategies under various time-varying constraints; and in [2] the optimization of delay tolerant bulk transfers over the Internet is formulated using space-time graphs.

III. Optimal Cost/Delay Tradeoff in DTNs

A. Cost/Delay Evolving Graphs

Let a set $\mathcal{N}$ of N nodes, forming a network whose directed communication link costs and storage costs are changing over time. In particular, time is divided in T **epochs**, each having a duration d^t, with $t = 1, 2, \ldots, T$. During each epoch, the set of directed links and their respective costs, as well as the storage costs, are fixed. Changes to the topology are instantaneous and happen during epoch transition times.

Clearly, this modeling of the network evolution is an approximation, as the properties of the network most often change smoothly and at arbitrary times. The model becomes more accurate as the number of epochs T increases and their durations d^t decrease. The advantage of the model is that it allows us to describe the evolution of the network in terms of a single graph, which we term **Cost/Delay Evolving Graph (C/DEG)**, due to its relation to the evolving graphs of [23].

C/DEGs are comprised of T directed subgraphs $\mathcal{G}^t = (\mathcal{N}^t, \mathcal{A}^t)$, $t = 1, \ldots, T$, called **replicas**, each related to a single epoch. Each vertex set $\mathcal{N}^t = \{1^t, 2^t, \ldots, N^t\}$ is a copy of the node set. The arc set $\mathcal{A}^t$ contains the **link arc**[1] (i^t, j^t) if node i can send data packets to node j during epoch t. Arc (i^t, j^t) is associated with a **link cost** c_{ij}^t modeling the cost of using the communication link from i to j during epoch t. We also connect each pair (i^t, i^{t+1}), $t = 1, \ldots, T-1$, $i = 1, \ldots, N$, with a **storage arc** of **storage cost** c_i^t. See Fig. 1 for an example C/DEG. Observe that a path on the C/DEG across multiple C/DEG vertices translates to a journey across time and space (i.e., nodes) on the original network.

B. Optimal Cost/Delay Curves

We define the **Optimal Cost/Delay Curve (OC/DC)** $C_{ij}(t)$ of the node pair (i, j), where $t = 1, \ldots, T$, as the function that gives the minimum cost with which node i can send a data packet, initially held by i at epoch 1, to node j, by epoch t *at the latest*. The minimum is taken over all paths connecting the vertex pairs (i^1, j^1), $(i^1, j^2), \ldots,$ (i^1, j^t) of the C/DEG. Therefore, each value $C_{ij}(t)$ is associated with an optimum path on the C/DEG, from node i^1 to one of the nodes j^1, $j^2, \ldots, j^t$, or equivalently an optimum journey that starts at i at epoch 1 and has ended at j by epoch t. Note that $C_{ii}(t) = 0$.

Due to the existence of storage costs, it is possible that the least-cost journey connecting nodes i and j with a delay of at most t has a delay strictly smaller than t. However,

[1]We use the convention that graphs are comprised of *arcs* and *vertices*, while networks are comprised of *nodes* and *(communication) links*.

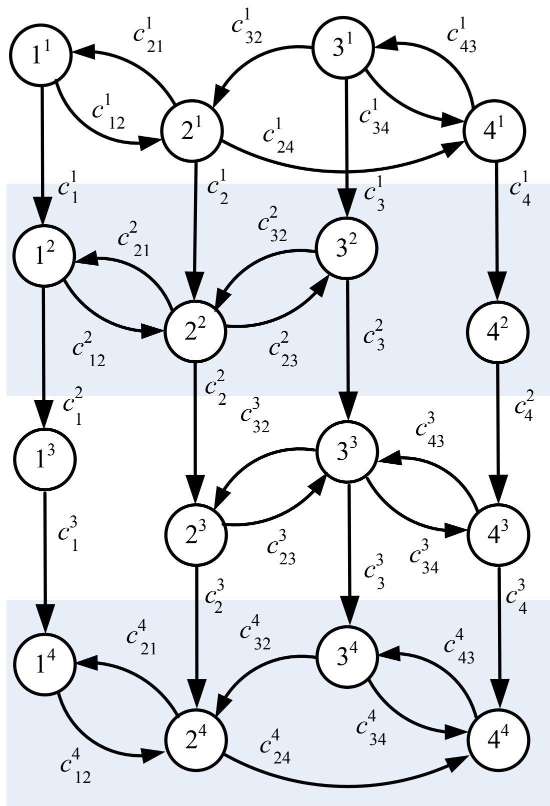

Figure 1. An example 4-replica, 4-node C/DEG.

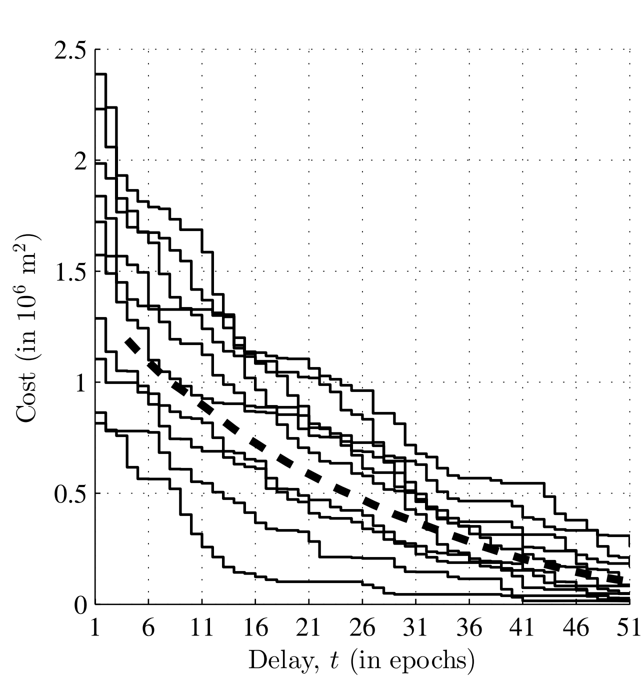

Figure 2. Ten sample OC/DCs (thin continuous lines) and the average of $50,000$ OC/DCs (thick dashed line), derived from 50 realizations of the 1001-node example network of Section III-C.

as $C_{ij}(t)$ is found by taking the minimum over all journeys of duration less than or equal to t, it follows that $C_{ij}(t)$ is a decreasing function of t. To emphasize this latter point, we also define the **Punctual Cost/Delay Curve (PC/DC)** $\widehat{C}_{ij}(t)$ of the node pair (i, j), where $i, j \in \mathcal{N}$ and $t = 1, \ldots, T$, as the minimum cost with which a data packet initially held by node i at epoch 1, can be at node j at *exactly* epoch t (and possibly having arrived earlier), where $t = 1, \ldots, T$. Therefore, the minimum is taken only over all paths connecting the vertex pair (i^1, j^t) of the C/DEG, and the PC/DCs $\widehat{C}_{ij}(t)$ need not be decreasing with time t.

C. Numerical Example

We consider $N = 1001$ nodes communicating over a common wireless channel. Node 1 is immobile, and acting as a Base Station (BS). All other nodes are constrained to move within a square of side $L = 10$ km, centered at the BS, for $T = 500$ epochs, each having a duration $d^t = 10$ sec. The mobile node locations during each epoch are those sampled, at the start of each epoch, from the following underlying smooth mobility model: each node moves towards a fixed direction with a constant speed $v = 36$ km/h; upon hitting a boundary, nodes get perfectly reflected; initially, nodes are placed independently of each other, uniformly in the square region, with a direction also chosen uniformly.

Any two nodes i, j within a distance d at most equal to a maximum communication range $R = 600$ m of each other can communicate directly, but with a cost equal to d^2. This quadratic cost penalizes long transmissions over shorter ones, and is a reasonable choice both for networks in which nodes have a limited energy supply, and therefore want to avoid transmissions over long distances, and also for networks where bandwidth is limited and therefore long transmissions should be avoided because they reserve bandwidth over a large area. For this numerical example, we set all storage costs to 0. Therefore, the transportation cost is only due to the transmissions, and is measured in m^2.

In Fig. 2 we plot 10 OC/DCs associated with 10 randomly chosen node pairs $(i, 1)$. We also plot the average of the $50,000$ OC/DCs corresponding to each pair $(i, 1)$, $i = 2, \ldots, 1001$, and for 50 independent network realizations.

For the value of the communication range R used, the network is almost always connected. However, there are a few pairs among the $50,000$ considered for which the packet delivery is not possible before 1 to 3 (depending on the case) epochs expire. Therefore, their OC/DCs are infinite for the first 1-3 epochs. As a result, the average of the $50,000$ OC/DCs is infinite (and hence not plotted) for epochs $1-3$.

D. Efficient Calculation of Optimal Cost/Delay Curves

In order to calculate the OC/DCs $C_{ij}(t)$, and the optimal C/DEG paths associated with them, we can first calculate the PC/DCs $\widehat{C}_{ij}(t)$ and the optimal C/DEG paths associated with those, and then find the OC/DCs $C_{ij}(t)$ using the fact that $C_{ij}(t) = \min\{\widehat{C}_{ij}(s),\ s = 1, \ldots, t\}$.

To calculate the $\widehat{C}_{ij}(t)$ and the optimal C/DEG paths associated with them, it is in principle possible, although not efficient, to execute any one-to-many shortest path algorithm on the complete C/DEG N times, once for every vertex i^1, $i = 1, \ldots, N$. The resulting computational complexity depends on the shortest-path algorithm used and its implementation details [27], however it is straightforward to show that the number of computations needed increase with T at least as fast as $T \log T$.

Fortunately, C/DEGs have two key features that permit the use of a modified recursive version of any one-to-many shortest path algorithm, in the process achieving a complexity which is always linear in T. In particular:

1) There are no arcs connecting a vertex of replica t to any vertex of replicas $1, \ldots, t-1, t+2, \ldots, T$.
2) The cross-replica storage arcs are all those, and only those, of the form (i^t, i^{t+1}), where $t = 1, \ldots, N-1$.

This structure implies that it is possible to execute a fast, custom-made shortest path algorithm on the whole C/DEG. The trick is to realize that one can determine the optimal paths appearing up to replica t, for $t = 1, \ldots, T$, without considering the links appearing in replicas t' for $t' > t$ (indeed, they cannot have an effect as no path entering a replica t' can visit any replica $t < t'$). In addition, the optimal paths ending at replica t are all extensions of some optimal path ending at replica $t-1$ (indeed, there are no links connecting non-consecutive replicas). Therefore, one can determine all optimal paths ending at nodes of replica 1, use these to determine the optimal paths ending at nodes of replica 2, and so one, while at any of the T steps the number of nodes and links involved are not a function of T, hence the complexity of Algorithm 1 is linear in T. Algorithm 1 formalizes this idea.

Algorithm 1: Fast calculation of PC/DCs, OC/DCs, and optimal C/DEG paths.

```
for i = 1 to N do
    Execute a shortest path algorithm on the first
    replica (N^1, A^1), with i^1 as the root;
    For all j = 1, ..., N, set Ĉ_ij(1) equal to the
    optimal cost from i^1 to j^1;
    for t = 2 to T do
        Create the virtual graph G^v = (N^t, A^t);
        Augment G^v by a virtual node i^0 and arcs
        connecting i^0 with all other nodes. The cost of
        (i^0, j^t) is set to Ĉ_ij(t-1) + c_j^{t-1};
        Execute a shortest path algorithm on G^v, with
        i^0 as the root;
        for j = 1 to N do
            if Shortest path from i^0 to j^t is (i^0, j^t) then
                Set the optimal path from i^1 to j^t by
                appending j^t to the optimal path from
                i^1 to j^{t-1}.
            else
                Let (i^0, k^t, P, j^t) be the shortest path
                from i^0 to j^t, for some replica t node
                k^t and sequence P of replica t nodes.
                Set the optimal path from i^1 to j^t by
                appending (k^t, P, j^t) to the optimal
                path from i^1 to k^{t-1}.
            end
            Set Ĉ_ij(t) equal to the optimal path cost
            from i^1 to j^t;
            Set C_ij(t) = min{Ĉ_ij(s), s = 1, ..., t}. If
            the minimum is achieved for s = k, then
            the optimal path associated with C_ij(t) is
            set to the optimal path from i^1 to j^k;
        end
    end
end
```

IV. Achievable Cost/Delay Tradeoffs in Wireless DTNs

We now focus on wireless DTNs using geographic routing. We first define a family of routing protocols with a common structure. We then define five members of this family and evaluate them in terms of the aggregate packet transmission cost they achieve as a function of the maximum permitted packet delay.

A. Protocol Family

Let R be the common **communication range** of the nodes. We also define the **restricted communication range** R' as a maximum hop length that packets are allowed to cover. R' is a tunable parameter at most equal to R and is used to explore different regions of the cost/delay tradeoff.

Consider a node A of the network. Let $\mathcal{G}_l^A$ be the **local graph** of A comprised of all the nodes within the communication range R of A, and all links connecting such nodes of length smaller than the restricted range R'. We assume that A has full information on $\mathcal{G}_l^A$. At the start of each epoch and for each potential destination D (or, if more practical, whenever A receives a packet for D), A executes a **Neighbor Evaluation Rule (NER)** (to be specified later) in order to find, among those nodes reachable through $\mathcal{G}_l^A$ (which we refer to as **candidate nodes**), the node B that is the best candidate for receiving packets destined for D. We refer to B as the **target node**. We stress that A counts itself among the candidates. Whenever A receives such packets during the epoch, it transmits them to node C that is the next hop on the minimum-cost path to B on $\mathcal{G}_l^A$. However, upon receiving such packets, C forwards them according to its own execution of the NER. The forwarding of a packet stops when

1) the packet arrives at D, or
2) a loop is detected, i.e., a node discovers that the packet has already passed, in the ongoing epoch, through a node C to which it is about to forward it, or
3) the NER specifies that the current holder is the best candidate for having the packet.

Regarding the second case, we note that the occasional appearance of loops is inescapable, even for well designed NERs, due to the distributed execution of the NER and the fact that neighboring nodes have different information about the network topology, and so may take contradictory routing decisions. In the latter two cases, the packet awaits for the epoch to expire and the topology to change, at which point the packet forwarding begins anew. We assume that the duration of each epoch is enough for all needed transmissions to take place.

B. Protocol Variants

Here, we discuss five protocol variants that differ on the precise NER specified. The first three (MOVE, AeroRP, MCpPR) are modifications of known forwarding rules; the last two (BRR, CR) are novel. Let A be a node carrying a packet, and D be that packet's destination.

MOVE: Node A picks as target that candidate, among those moving towards D (i.e., whose distance to D decreases with time), that is scheduled to pass the closest to D, assuming that all candidates maintain their current velocity vector. If all candidates move away from D, then A picks the one closest to D.

AeroRP: Node A selects as target that candidate moving towards D with the smallest TTI, as this is calculated in (1). If no candidate moves towards D, A keeps the packet.

Min-Cost-per-Progress Rule (MCpPR): For each candidate B closer to D than A, let the **cost/progress ratio**[2]

$$r'_{AB} = \frac{C_{A\to B}}{|AD| - |BD|}, \tag{2}$$

where $C_{A\to B}$ is the cost of sending the packet from A to B (as discussed, using a minimum cost path on $\mathcal{G}_l^A$). All other candidates have a cost/progress ratio equal to ∞. Node A selects as target the candidate with the smallest cost/progress ratio, if this is finite; otherwise A keeps the packet.

Balanced Ratio Rule (BRR): For each candidate B moving towards D and for which $|AD| - |DZ| > 0$, where Z is the point where B will be closest to D according to its current velocity vector, let the following **balanced ratio**

$$r''_{AB} = \frac{C_{A\to B} + a d_{B\to Z}}{|AD| - |DZ|}, \tag{3}$$

where $d_{B\to Z}$ is the delay for B to arrive at Z and a is a positive real coefficient termed the **conversion coefficient**.

[2]We denote by $|XY|$ the distance between points X and Y.

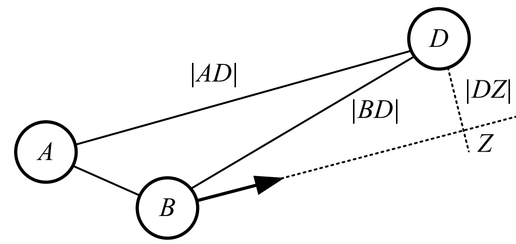

Figure 3. Topology used in the definition of MCpPR, BRR, and CR.

For all other candidates, the balanced ratio is set to ∞. If there are candidates with finite balanced ratios, node A selects as target the one with the smallest balanced ratio, otherwise it keeps the packet.

BRR may be thought of as a non-greedy version of MCpPR: instead of trying to maximize *immediate* gains when selecting the next hop, BRR focuses instead on the gains made by physically transporting the packet, taking into account the *long-term* progress made, the delay to materialize this progress, and the transmission cost that must be invested initially. The parameter a is used to strike a balance (i.e., between keeping both the cost and the delay small) that is suitable for the application that originated the packet.

Composite Rule (CR): Node A calculates the following **composite metric**:

$$c_{\mathrm{AB}} = \min\left\{r'_{AB}, r''_{AB}\right\},$$

where r'_{AB} and r''_{AB} are calculated in (2), (3), respectively, and selects as target the candidate with the smallest composite metric. If there is no candidate with a finite composite metric, A keeps the packet.

The rationale behind the use of the composite metric is that we would like from a node holding a packet to be ready to take any opportunity arising, and be ready to employ either low cost hops with *immediate* gains in the progress made to the destination, or hops that *eventually* lead to a significant reduction to the distance to the destination with an attractive combination of cost and delay.

C. Achievable Cost/Delay Curves

We will evaluate the performance of all protocol variants in terms of the **Achievable Cost/Delay Curve (AC/DC)** of each, and its relation to the respective OC/DC.

For each node pair (i, j), and each protocol X, the AC/DC $C_{ij}^X(t)$ is defined as the curve that gives, for each t, the minimum aggregate transport cost that this protocol can achieve for the pair in question, with a delay of at most t. If the protocol cannot deliver the packet with delay at most t, then $C_{ij}^X(t) = \infty$.

If the protocol X has tunable parameters, then this minimum transportation cost is achieved minimizing over the complete range(s) of the tunable parameter(s) that lead to

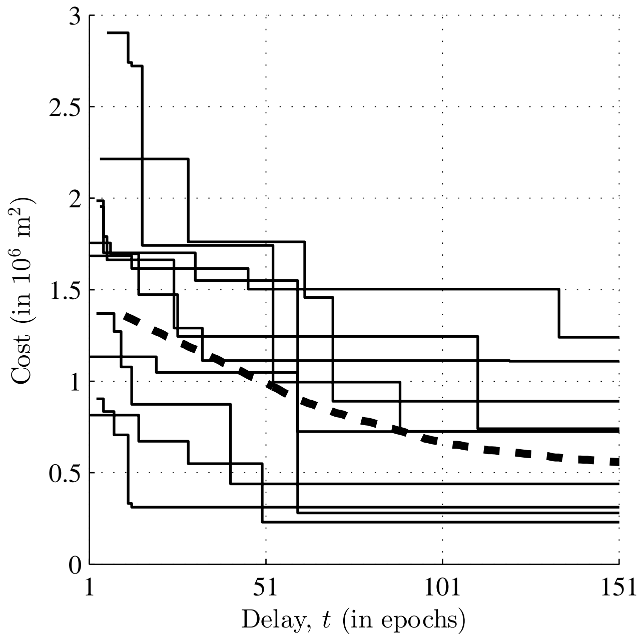

Figure 4. Ten sample AC/DCs derived using MCpPR (thin continuous lines) and the average of $50,000$ AC/DCs (thick dashed line), for the example network of Section III-C.

packet delays of at most t. If there are no tunable parameters, then either the protocol cannot deliver the packet from i to j with any delay, in which case $C_{ij}^X(t) = \infty$ for all t, or it delivers it by time t_0, in which case $C_{ij}^X(t) = \infty$ for $t < t_0$ and $C_{ij}^X(t) = k$ for $k \geq t_0$.

Here, we limit the usage of AC/DCs in the discrete-time network model case where time is measured in epochs and the network changes abruptly only during epoch transitions. We note, however, that AC/DCs are applicable in the more common case where time flows continuously.

As an example, we plot in Fig. 4 the AC/DCs of the ten sample node pairs $(i, 1)$ of Fig. 2, as well as the average of all 1000 curves for each node $i = 2, \ldots, 1001$, and for the 50 independent realizations of the topology of Fig. 2. The protocol used is MCpPR, and each curve is traced by simulating the routing of the packet for 10 different values of R'[3]. This gives 10 points on the cost/delay plane which in turn can be used to estimate the corresponding curve. We note that some of these 10 points on the cost/delay plane will coincide, and that there may be multiple distinct points with the same delay, in which case only the point with the smallest cost will have an effect on the curve.

[3]It was found experimentally that increasing this modest number of values does not affect the results significantly.

D. Performance Evaluation

First, we revisit the 1001-node network of Section III-C maintaining the same mobility model. However, regarding the velocity, we now consider two cases:

1) Node speeds are equal to 36 $\mathrm{km/h}$, as previously.
2) Nodes speeds are random, independent, and uniformly selected in the range from 0 to 36 $\mathrm{km/h}$.

For each of these cases, we consider two different models for the communication range R and transmission cost $c(d)$ over distance d:

1) $R = 600$ m and $c(d) = d^2$, as previously.
2) $R = 300$ m and $c(d) = R^2 = 9 \times 10^5$ m^2.

Therefore, we create four cases that collectively cover a relatively wide range of scenarios regarding the levels of connectivity, the speed profiles, and the nature of the costs.

In Fig. 5 we plot, for each case, the averages of the OC/DC and the five AC/DCs of pairs $(i, 1)$, $i = 2, \ldots, 1001$, and for 50 realizations of the network. Therefore, each curve represents an average of $50,000$ curves, thus the statistical error is very small. For BRR and CR, the curves were found considering, in addition to the 10 values of R', 45 values of a ranging from 0 to 10^6 m^2/epoch.

Note that, for many node pairs $(i, 1)$ and choices of the parameters, it is impossible for the packet to be delivered, either over the optimal path or with a routing protocol, with a delay less than some minimum which is greater than 1, and so the value of the corresponding curve is ∞. In drawing the averages, we do not count these infinite values, but refrain from drawing the average curve if more than 1% of the values equal ∞. Therefore, the range over which the average curve is plotted is an important indication about the performance of the corresponding protocol.

We also consider two interesting scenarios where the node movements are constrained, through different mechanisms. All the previously applied assumptions hold, with the following modifications, for each case:

In the **home region setting**, each node moves as in Section III-C but within a **home region** square of side $L' = L/5 = 2$ km. Home region squares are chosen so that they always completely fall within the larger square of side L, but otherwise uniformly, and independently of each other. (Note that this model is somehow reminiscent of the mobility model used in [28].)

In the **urban setting**, nodes are constrained to move within a square of side $L = 10$ km, but along vertical and horizontal highways, separated from each other by a distance of 100 m. Nodes move according to the random waypoint model, and travel from a point A to a point B using any of the shortest paths available chosen randomly. In this case, when a protocol needs to calculate the distance between a point and a curve, it is using the Manhattan distance. In (1) the *average* relative velocity s_d is calculated using the current location of the node, its destination, and the time to arrive there (as opposed to using its *instantaneous* relative velocity, which is a poor indicator of the node's real movement).

In Fig. 6 we plot, for each setting, the averages of the OC/DC and the five AC/DCs of pairs $(i, 1)$, $i = 2, \ldots, 1001$, and for 50 realizations of the network. The speed of nodes is equal to $36\ \mathrm{km/h}$, and we assume that $R = 600\ \mathrm{m}$ and $c(d) = d^2$. For BRR and CR, the curves were found considering the values of R' and a used in Fig. 5.

E. Discussion

A number of observations are in order. Firstly, for the settings considered here, CR has a clear advantage over the other variants: in the case of the quadratic cost, the average curve of CR is the closest to the average OC/DC, for the complete range of delays. However, in the two cases of constant cost, although CR consistently delivers more packets for smaller delays, using BRR and tolerating large delays leads to lower costs.

Observe that in the case of constant cost, due to the smaller R used, the probability that a packet can be delivered by epoch t drops below 1% for large values of t, especially for the case of the uniformly distributed speeds.

Another interesting result is that CR always behaves much better than both BRR and MCpPR, as opposed to performing as well as the best of the two, although it is no more than a composition of the two. This remarkable improvement is due to the adaptability of CR: a packet following CR will propagate towards its destination by alternatively using BRR and MCpPR, using at any given time the most favorable protocol of the two.

Also, the performance of MOVE is comparable to that of AeroRP, when all nodes move with the same speed. If, however, nodes move with variable speeds, AeroRP is clearly superior. In fact, for the case of variable speed and fixed cost, the probability that the AC/DC of MOVE is finite drops below 1% for $t \geq 400$. Therefore, that curve does not appear in the plot.

Another interesting result is the poor performance of MCpPR in terms of the cost/delay tradeoff it achieves. As MCpPR is the only rule that does not take into account the speed relative to the destination of potential next hops, it follows that this relative speed is a key parameter.

The performance of all protocols is farthest from the optimum in the case of the home region setting. The explanation for this is the fact that int he home region setting nodes change their direction frequently, and therefore the current direction of a node is not such a great indicator about how useful it will turn out to be.

Although our performance evaluation is preliminary, it is interesting to note that that BRR and CR perform better than MOVE, AeroRP, and MCpPR, with BRR having a slight advantage over CR in terms of the cost it achieves in the large-delay range, and the CR having a clear advantage over BRR in the low-delay range.

As a final note, a closer study of our simulation results shows how nodes can tune the parameters available to them in order to trace the cost/delay tradeoff achievable by the rules they adopt. In the cases of MCpPR, MOVE, and AeroRP, reducing R' invariably leads to smaller costs but greater delays for the packets. In the cases of BR and CRR, reducing either R' or a leads to smaller costs and greater delays.

V. Conclusions

Our work makes two immediate contributions: Firstly, we present a formulation, in terms of C/DEGs and OC/DCs, for studying the fundamental tradeoff between the maximum permitted packet delivery delay and the minimum possible aggregate transportation cost (that includes both storage and transmission costs) existing in *all* DTNs.

Secondly, we study this tradeoff specifically in the context of geographically routed wireless DTNs, and evaluate the performance of a number of known and novel forwarding rules in terms of the cost/delay tradeoff they achieve. Of the two forwarding rules we propose, CR exhibits a cost/delay tradeoff closest to the optimal for a variety of scenarios, while BRR achieves the lowest costs for large delays and a fixed cost model.

A final, implicit, contribution of this work, and perhaps the most important, is that we set the agenda for a systematic study of cost/delay tradeoffs in a variety of DTN settings. For example, an important problem raised in our work in the context of geographical routing is finding, among all forwarding rules that do not make use of information on the future topology, which one is the 'best' performing in terms of the gap between its AC/DC and the OC/DC. Similar questions can be formed in other DTN settings.

References

[1] S. Jain, K. Fall, and R. Patra, "Routing in a delay tolerant network," in *Proc. ACM SIGCOMM*, Portland, OR, Aug.-Sep. 2004, pp. 145–157.

[2] N. Laoutaris, G. Smaragdakis, P. Rodriguez, and R. Sundaram, "Delay tolerant bulk data transfers on the internet," in *Proc. ACM Sigmetrics 2009*, Seattle, WA, June 2009, pp. 229–238.

[3] T. Small and Z. J. Haas, "The shared wireless infostation model – a new ad hoc networking paradigm (or where there is a whale, there is a way)," in *Proc. ACM MOBIHOC*, Annapolis, MD, June 2003.

[4] M. Grossglauser and D. Tse, "Mobility increases the capacity of ad-hoc wireless networks," in *Proc. IEEE INFOCOM*, vol. 3, Anchorage, AL, Apr. 2001, pp. 1360–1369.

[5] S. Toumpis and A. J. Goldsmith, "Large wireless networks under fading, mobility, and delay constraints," in *Proc. IEEE INFOCOM*, Hong Kong, China, Mar.-Apr. 2004.

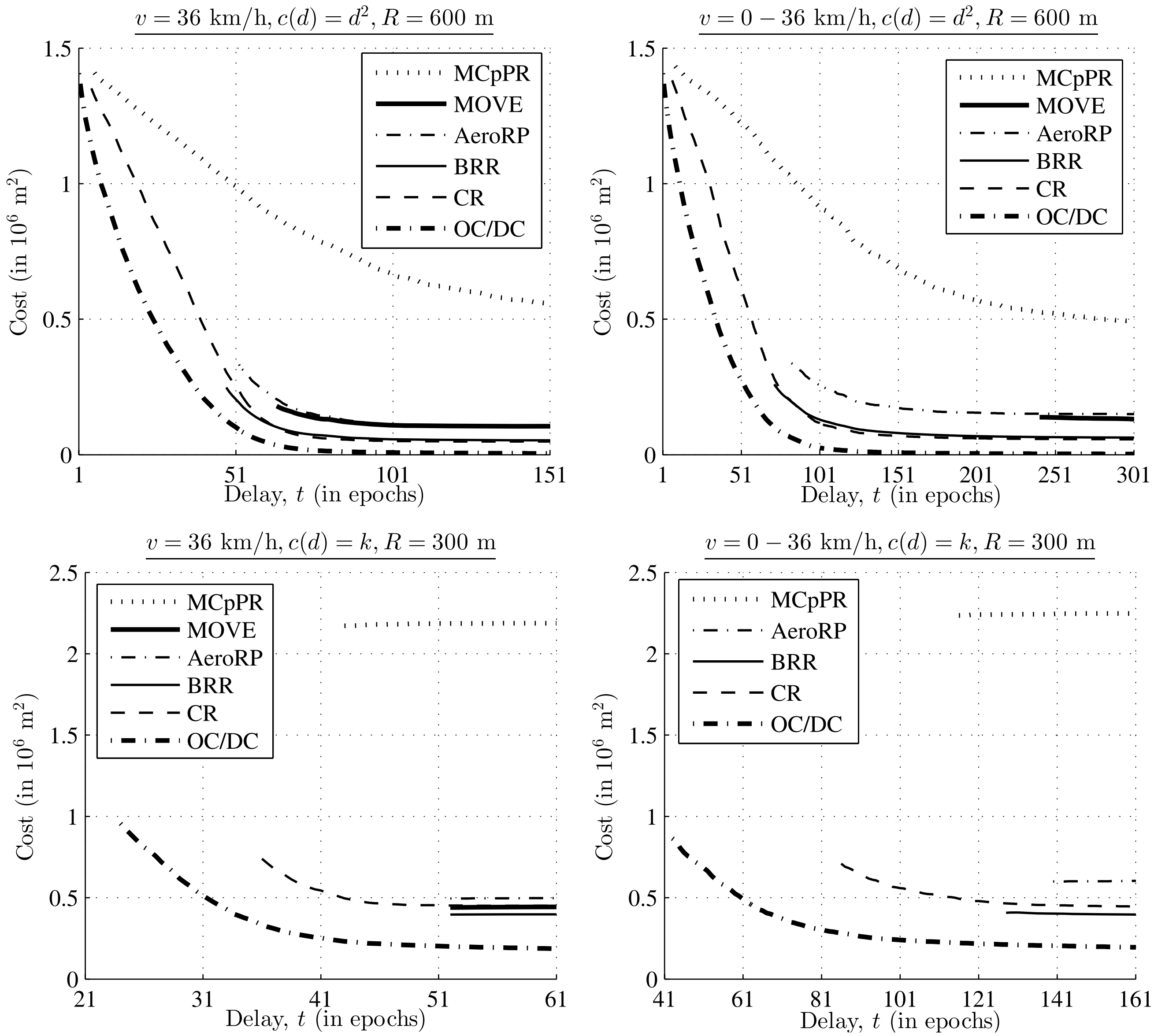

Figure 5. Averages of optimal cost/delay curves and achievable cost/delay curves for five different protocol variants: MOVE, AeroRP, the Min-Cost-per-Progress Rule, the Balanced Ratio Rule, and the Composite Rule. Two choices for the node speed are considered: all nodes move with constant speed $v = 36$ km/h, and each node moves with speed uniformly chosen in the $v = 0 - 36$ km/h range. Two combinations of communication range R and cost models $c(d)$ are also considered: $R = 600$ m, $c(d) = d^2$ and $R = 300$ m, $c(d) = R^2 = 9 \times 10^5$ m^2. All curves are the averages of $50,000$ curves corresponding to specific packets.

[6] P. Juang, H. Oki, Y. Wang, and M. Martonosi, "Energy-efficient computing for wildlife tracking: Design tradeoffs and early experiments with zebranet," in *Proc. ASPLOS-X*, San Jose, CA, Oct. 2002.

[7] R. C. Shah, S. Roy, S. Jain, and W. Brunette, "Data MULEs: Modeling and analysis of a three-tier system for sparse sensor networks," *Ad Hoc Networks*, vol. 1, no. 2-3, pp. 215–233, Sep. 2003.

[8] E. Kuiper and S. Nadjm-Tehrani, "Geographical routing with location service in intermittently connected MANETs," *IEEE Trans. on Veh. Technol.*, vol. 60, no. 2, pp. 592–604, Feb. 2011.

[9] P.-C. Cheng, K. Lee, M. Gerla, and J. Härri, "GeoDTN+Nav: Geographical DTN routing with navigator prediction for urban vehicular environment," *Mobile Networks and Applications*, vol. 15, no. 1, pp. 61–82, Apr. 2010.

[10] A. Sidera and S. Toumpis, "DTFR: A geographic routing protocol for wireless delay tolerant networks," in *Proc. Med-Hoc-Net*, Favignana Island, Italy, June 2011.

[11] R. Nelson and L. Kleinrock, "The spatial capacity of a slotted

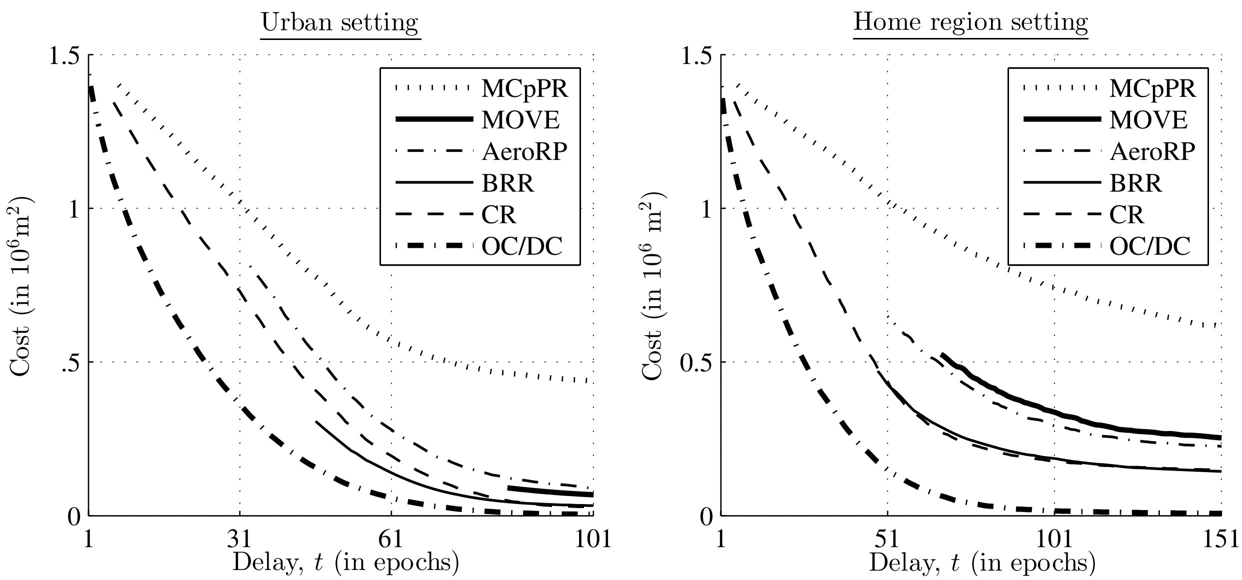

Figure 6. Averages of optimal cost/delay curves and achievable cost/delay curves for five different protocol variants of Fig. 5 and for two settings: the urban setting and the home region setting. All nodes move with constant speed $v = 36$ km/h. Also, $R = 600$ m, $c(d) = d^2$ and $R = 300$ m, $c(d) = R^2 = 9 \times 10^5$ m^2. All curves are the averages of $50,000$ curves corresponding to specific packets.

ALOHA multihop packet radio network with capture," *IEEE Trans. on Commun.*, vol. COM-32, no. 6, pp. 684–694, June 1984.

[12] S. Basagni, I. Chlamtac, V. R. Syrotiuk, and B. A. Woodward, "A distance routing effect algorithm for mobility," in *Proc. ACM MobiCOM*, Dallas, TX, Oct. 1998, pp. 76–84.

[13] B. Karp and H.-T. Kung, "GPSR: greedy perimeter stateless routing for wireless networks," Boston, MA, Aug. 2000.

[14] M. Mauve, A. Widmer, and H. Hartenstein, "A survey on position-based routing in mobile ad hoc networks," *IEEE Network*, vol. 15, no. 6, pp. 30–39.

[15] K. A. Harras and K. Almeroth, "Controlled flooding in disconnected sparse mobile networks," *Wireless Communications and Mobile Computing (WCMC) Journal*, vol. 9, no. 1, pp. 21–23, Jan. 2009.

[16] D. G. J. LeBrun, C.-N. Chiah and M. Zhang, "Knowledge-based opportunistic forwarding in vehicular wireless ad hoc networks," in *Proc. IEEE VTC Spring*, vol. 4, Florence, Italy, May-June 2005, pp. 2289–2293.

[17] K. Peters, A. Jabbar, E. K. Çetinkaya, and J. P. G. Sterbenz, "A geographical routing protocol for highly-dynamic aeronautical networks," in *Proc. IEEE WCNC*, Cancun, Mexico, Mar. 2011.

[18] I. Leontiadis and C. Mascolo, "GeOpps: Geographical opportunistic routing for vehicular networks," in *Proc. IEEE WOWMOM*, Helsinki, Finland, June 2007.

[19] L. R. Ford, Jr. and D. R. Fulkerson, "Constructing maximal dynamic flows from static flows," *Operations Research*, vol. 6, no. 3, pp. 419–433, May-June 1958.

[20] W. B. Powell, P. Jaillet, and A. Odoni, "Stochastic and dynamic networks and routing," *Handbooks in Operations Research and Management Science*, vol. 8, pp. 141–295, 1995.

[21] B. Kotnyek, "An annotated overview of dynamic network flows," INRIA, Tech. Rep. RR-4936, 2003 2003.

[22] M. Skutella, "An introduction to network flows over time," *Springer Research Trends in Combinatorial Optimization*, pp. 451–482, 2009.

[23] A. Ferreira, A. Goldman, and J. Monteiro, "Performance evaluation of routing protocols for MANETs with known connectivity patterns using evolving graphs," *Wireless Networks*, vol. 16, no. 3, pp. 627–640, Apr. 2010.

[24] A. Ferreira and A. Jarry, "Complexity of minimum spanning tree in evolving graphs and the minimum-energy broadcast routing problem," in *Proc. WiOpt*, Cambridge, UK, Mar. 2004.

[25] S. Merugu, M. Ammar, and E. Zegura, "Routing in space and time in networks with predictable mobility," Georgia Institute of Technology, Tech. Rep. GIT-CC-04-07, 2004, available at `http://hdl.handle.net/1853/6492`.

[26] D. Hay and P. Giaccone, "Optimal routing and scheduling for deterministic delay tolerant networks," in *Proc. IEEE WONS*, Snowbird, UT, Feb. 2009.

[27] T. H. Cormen, C. E. Leicerson, R. L. Rivest, and C. Stein, Eds., *Introduction to Algorithms*, 2nd ed. MIT Press, 2001.

[28] K. A. Harras and K. Almeroth, "Scheduling messengers in disconnected clustered mobile networks," *Ad Hoc and Sensor Wireless Networks*, vol. 9, no. 3-4, pp. 275–304, 2010.