Impact of Correlated Mobility on Delay-Throughput Performance in Mobile Ad-Hoc Networks

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

Impact of Correlated Mobility on Delay-Throughput Performance in Mobile Ad-Hoc Networks

Delia Ciullo *, Valentina Martina *, Michele Garetto †, Emilio Leonardi *
* Dipartimento di Elettronica, Politecnico di Torino, Torino, Italy
† Dipartimento di Informatica, Università di Torino, Torino, Italy

***Abstract*—We extend the analysis of the scaling laws of wireless ad hoc networks to the case of correlated nodes movements, which are commonly found in real mobility processes. We consider a simple version of the Reference Point Group Mobility model, in which nodes belonging to the same group are constrained to lie in a disc area, whose center moves uniformly across the network according to the i.i.d. model. We assume fast mobility conditions, and take as primary goal the maximization of per-node throughput. We discover that correlated node movements have huge impact on asymptotic throughput and delay, and can sometimes lead to better performance than the one achievable under independent nodes movements.**

I. Introduction and related work

In the last few years the *store-carry-forward* communication paradigm, which allows nodes to physically carry buffered data as they move around the network area, has opened an entire new area of research with many promising applications in the context of delay-tolerant networking [1].

In their seminal work [2], Grossglauser and Tse have shown that mobile nodes employing the *store-carry-forward* paradigm can achieve constant throughput even when the number of nodes grows to infinity, in contrast to the severe throughput decay (like $1/\sqrt{n}$) incurred in fixed networks [3]. The basic requirement of their 2-hop scheme is that nodes uniformly visit the entire network space according to an arbitrary, stationary and ergodic mobility process with *independent* trajectories.

When considering also the delay performance, the specific details about how nodes move become important. Several papers have analyzed throughput-delay trade-offs for various mobility models, ranging from the simple reshuffling model (also referred to as i.i.d. model) [4], [5], [6], to the Brownian motion [7], and variants of random walk and random waypoint [8], [9]. The impact of limited buffers has been considered in [10]. In all these works, the mobility of the nodes has always been assumed to be uncorrelated (i.e. independent from node to node) and uniform over the area.

Some works have already considered the impact on the capacity of restricted mobility models (i.e. relaxing the assumption that nodes uniformly visit the network area) [11], [12], [13], still maintaining the independence assumption on the nodes mobility process.

To the best of our knowledge, no work has been done so far to investigate the impact of correlation among nodes movements on the asymptotic throughput and delay of large mobile networks. This is rather surprising in light of the fact that real mobility processes (of pedestrians, vehicles, animals) exhibit significant degrees of correlation, as observed in several traces [14], [15], [16], [17].

The goal of our work is to study, for the first time, the scaling laws of capacity and delay for large mobile networks including correlated nodes movements. To this aim, we consider a very simple model of correlated mobility based on the popular Reference Point Group Mobility (RPGM) model introduced in [18]. Nodes are organized into several groups, and the mobility of nodes belonging to the same group is confined within a disc area. Each group has a logical center, which moves around the network according to the i.i.d. mobility model, dragging behind all nodes belonging to it. Notice that in the long run each node visits uniformly the entire network space, however the trajectories of individual nodes are not independent because they are constrained to jointly follow their respective groups. By changing a few parameters, our model allows to explore various degrees of correlation in the node mobility process.

We propose novel scheduling-routing schemes whose primary goal is to maximize the per-node throughput. As a secondary goal, we also seek to minimize the packet delivery delay. Our main finding is that node correlation has a strong impact on both throughput and delay performance. Interestingly, correlated mobility can lead both to better and to worse performance with respect to the case in which node movements are independent.

Prior to our work, the impact of correlated node movements on existing and novel routing protocols has been extensively investigated by simulation. In the context of traditional *store-and-forward* networks, [19] analyzed the effect of various mobility models, including correlated movements, on classical routing protocols (DSR, AODV), while in [20] the authors have proposed a novel routing protocol, called LANMAR, which directly exploit group mobility patterns to improve routing efficiency. Similarly to our scheme, they propose a hierarchical approach in which data are first routed at the group level, and then routed within the group containing the destination. A similar idea is proposed in [21] for the *store-carry-forward* communication paradigm. In particular, a history-based approach similar to PRoPHET [22] is adopted at the group level. Like us, the authors of [21] also employ a replication strategy to improve the delivery delay. We emphasize that previous work relied entirely on simulations to evaluate the performance of the proposed schemes, without analyzing asymptotic scaling laws nor the optimality of the proposed solutions in terms of system throughput and delay.

978-1-4244-5837-0/10/$26.00 ©2010 IEEE

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

II. System assumptions

A. Mobility Model

We consider an extended network comprising n nodes moving over a square region $\mathcal{O}$ of area n with wrap-around conditions (i.e., a torus), to avoid border effects. Note that, under this assumption, the overall node density over the area remains constant and equal to 1, as we increase n.

We assume that nodes are partitioned into m groups, with $m = \Theta(n^\nu)$, $\nu \in [0, 1)$. For simplicity, we assume that each group comprises an integer number $q = n/m$ of nodes. Note, however, that our results would not change, in scaling order, if the cardinality of the groups were not exactly the same, as long as each group contains $\Theta(n/m) = \Theta(n^{1-\nu})$ nodes.

Time is divided into slots of equal duration, which is normalized to 1. Nodes belonging to the same group move over the network area in a correlated fashion. To model this behavior, we assume that, at any given slot, all nodes of a group have to reside concurrently within a same portion, of area $o(n)$, of the total network space. In the following we will refer to such a portion as the *cluster-region* or simply the *cluster*, associated to the group.

We assume that each cluster-region has a circular shape of radius R. We can explore various degrees of correlation in the node mobility process by letting R scale with n as well, as $R = \Theta(n^\beta)$, with $\beta \in [0, 1/2)$. Notice that $\beta = 0$ corresponds to the extreme case in which each group occupies a constant fraction of the network area (just as if all nodes of a group were located at a single point), irrespective of the number of nodes in it.

We have yet to specify how nodes actually move over the network area from one slot to another. The mobility process of a given node i belonging to group j is described by the combination of two movements: i) a group movement (i.e., the shift of the cluster-region associated to group j during a slot); ii) a node movement (i.e., the change of position of node i within the cluster-region of group j).

For what concerns the group movement, we assume that each cluster-region has a center point, whose position is updated at each time slot by choosing a new location uniformly at random in the network area, independently for each group. This is similar to the so-called reshuffling model, or bi-dimensional i.i.d. mobility model, considered in previous work [4], [5], [6], however here we adopt this model only to update the positions of the cluster centres. The mobility processes of individual nodes are not independent in our model because, once the new position of a cluster centre has been selected, all nodes belonging to the corresponding group have to move to a place close to it (i.e., inside a region of area $R^2 = o(n)$ around the cluster centre). We observe that the degree of correlation in the node mobility process increases as we either i) reduce the area of each cluster-region (smaller values of β); ii) reduce the number of groups (smaller values of ν).

For what concerns the movement of nodes within their cluster-regions, we consider two extreme cases i) the *reshuffling* model, according to which each node, independently of others, in the next slot moves to a position chosen uniformly at random in the cluster-region; ii) the *crystallized* model, in which nodes belonging to the same group maintain their relative positions within the cluster-region indefinitely, starting from an initial configuration in which they are placed uniformly at random in the cluster-region. More realistic models of node movements inside their cluster-region would produce a mobility degree in between the two cases above (as suggested in [23]), hence our model allows to identify the feasible range of system performance as we vary the mobility degree of nodes within a cluster-region.

B. Communication Model

To account for interference among simultaneous transmissions, we adopt the protocol model introduced in [3] and widely used in the literature[1]. According to the protocol model, nodes employ a common range r for all transmissions which occur in the same time slot (r can be different from slot to slot); equivalently, they employ a common power level in each slot. A transmission from node i to node j using transmission range r can be successfully received at node j if and only if the following two conditions hold:

1) the distance between i and j is smaller than or equal to r, i.e., $d_{ij}(t) \leq r$.
2) for every other node k simultaneously transmitting, $d_{kj}(t) \geq (1 + \Delta)\, r$, being Δ a guard factor.

Transmissions occur at fixed rate which is normalized to 1. Moreover, we consider fast mobility conditions, according to which data can be transmitted over just one hop during any slot[2].

C. Traffic Model

Similarly to previous work we consider permutation traffic patterns in which every node is origin and destination of a single traffic flow of rate λ. Hence there are n source-destination (S-D) pairs in the network.

III. Summary of results and paper organization

Recall that the primary goal of our schemes is to maximize the system throughput. As a secondary goal, we seek to minimize delay. Hence we do not explore the full range of possible capacity-delay trade-offs. However, we are able to prove the optimality of our schemes under the considered goals, and we can conclude that correlated mobility can have a remarkable positive impact on network performance.

Depending on the values of β and ν two different regimes are possible. When $\nu + 2\beta < 1$ the sum of all cluster areas mR^2 is $o(n)$. This means that at any time clusters cover only a negligible fraction of the entire network area. Actually they form several small, disconnected and highly dense regions (the node density within a cluster is $\frac{n}{mR^2} = \omega(1)$) floating over a huge empty space. Spatial overlaps between different clusters are sporadic. We refer to this regime as the *cluster sparse* regime. In this case, the optimal per-node throughput, which can be achieved by our scheme both in the *reshuffling* and in the *crystallized* model, is $\lambda = \Theta(mR^2/n)$. Concerning the delay, in the *reshuffling* model

[1]Our results would not change under the physical model defined in [3], provided that the power loss exponent is larger than 2.

[2]We leave to future work the extension of the analysis to the slow mobility case, in which multi-hop transmissions can be performed during the same slot.

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

$D = \Theta\left(\max\left\{\frac{n}{R^2}, \frac{mR^4}{n}\right\}\right)$, whereas in the *crystallized* model $D = \Theta\left(\max\left\{\frac{n}{R^2}, mR^2 \log n\right\}\right)$. We emphasize that in all cases our schemes always make use of transmission ranges $O(1)$; thus they do not need to scale up the transmission power as n increases.

When $\nu + 2\beta > 1$ the sum of all cluster areas mR^2 is $\omega(n)$. This means that clusters cover the entire network area. Moreover, they are largely overlapped. We refer to this regime as the *cluster dense* regime. In this case we get $\lambda = \Theta(1)$ and $D = \Theta(n)$, like in the original Grossglauser-Tse scenario with independent node movements.

Figures 1 and 2 provide a graphical representation of our results on the throughput-delay plane for the *reshuffling* and the *crystallized* model, respectively. Shaded regions denote all optimal operating points that are obtained as we vary the parameters β and ν of our model. The grey scale is related to the degree of correlations in the mobility process: higher correlation, which results from reducing either β or ν, corresponds to lighter grey.

On both plots we have reported the line $D = n\lambda^2$ which denotes (neglecting logarithmic factors) the best possible throughput-delay trade-offs that can be obtained under independent reshuffling of all nodes (and fast-mobility conditions), according to [5], [6]. [3]

We observe that in the *reshuffling* model there are points above the line $D = n\lambda^2$. This means that in our correlated mobility model it is possible to obtain significant better performance than that achievable under uncorrelated mobility. On the contrary, all points in the *crystallized* model are below the line $D = n\lambda^2$. We have also reported on the plots a few curves that are obtained when we fix one of the parameters of the model (either β or ν), letting the other vary. We observe a quite complex range of possible behavior, especially in the *reshuffling* model. For large values of β and/or ν, the system operates in the *cluster-dense* regime, at the point $\lambda = \Theta(1)$ and $D = \Theta(n)$. As we increase the degree of correlation, by reducing either β or ν, at some point the system shift to the *cluster-sparse* regime. Here, in the case of the *reshuffling* model, significant better delays can be obtained with just a little penalty in system throughput. The best operating point (marked with P in Figure 1), can be approached when both $\beta \to 1/4$ and $\nu \to 1/2$. The introduction of additional correlation in the node mobility process (smaller values of β and ν) does not help, and eventually brings the system below the line $D = n\lambda^2$.

The rest of the paper is organized as follows. We will first consider the *cluster sparse* regime, which is more interesting and challenging to analyze. Indeed, the study of the *cluster dense* regime is very simple, since here we can apply the same strategies developed for nodes uniformly visiting the entire network space according to i.i.d. patterns. For the *cluster sparse* case, we will separately consider the *reshuffling* model in Section IV, and the *crystallized* model in Section V. The *cluster dense* regime will be briefly discussed in Section VI for both the *reshuffling* and the *crystallized* model[4]. We conclude in Section VII.

[3]Observe that the law $D = n\lambda^2$ is achievable for the i.i.d. reshuffling model only when the transmission range is allowed to scale to infinite as n increases; $D = n\lambda$ (neglecting poly-log terms) provides the best achievable trade-offs when the transmission range is constrained to be $O(1)$.

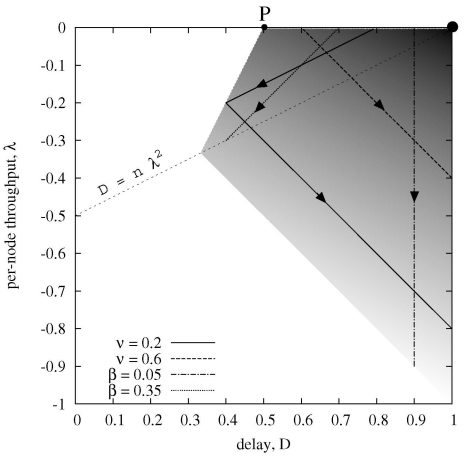

Fig. 1. Throughput-delay scaling for the *reshuffling* model. The marks on the axes represent the orders asymptotically in n.

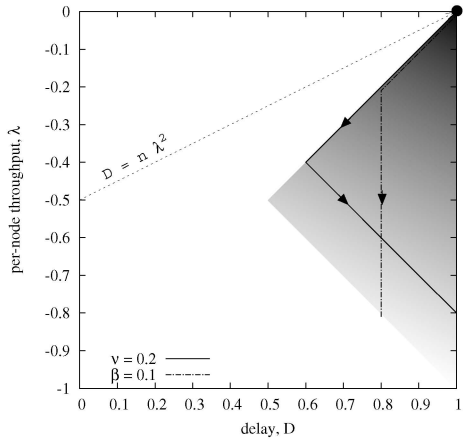

Fig. 2. Throughput-delay scaling for the *crystallized* model. The marks on the axes represent the orders asymptotically in n.

IV. Cluster sparse regime: reshuffling model

We will first introduce the scheduling-routing scheme that we have developed for this case, describing in particular the routing scheme in Section IV-A and the associated scheduling scheme in Section IV-B. Then in Section IV-C we will analyze the performance of the proposed scheme and prove its optimality.

A. Routing scheme

We propose a multi-hop routing scheme that generalizes the 2-hop scheme introduced by Grossglauser and Tse [2].

We focus on a particular traffic stream $s \to d$. Let C_s denote the cluster containing s (i.e., $s \in C_s$) and C_d the cluster containing d (i.e., $d \in C_d$). We neglect the particular case in which $C_s = C_d$, since w.h.p. s and d belong to different clusters (this is also the more stressful case for the system).

The rationale of our routing scheme is to first reach a node within the destination cluster C_d in the most efficient way,

[4]We leave for future investigations the special case in which $\nu + 2\beta = 1$

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

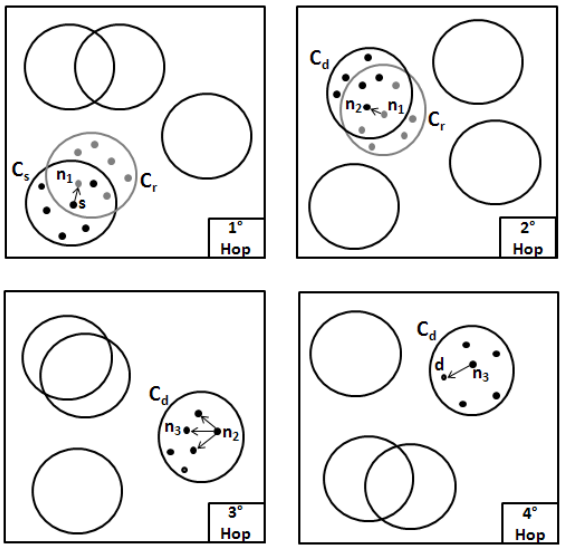

Fig. 3. Illustration of the 4-hop routing scheme

and then to forward the packets[5] within C_d up to the final destination d. We anticipate that the system throughput is bottlenecked in the first phase of the route, in which data has to reach the destination cluster. This is due to the fact that close contacts among nodes belonging to different clusters are rare, since they occur only when two clusters overlap in space.

The same principles that inspired the 2-hop scheme of Grossglauser and Tse suggest that the most efficient way to bring a message within the destination cluster is to adopt a 2-hop relaying scheme (at the cluster level), in which each packet transits through a random intermediate cluster C_r. This allows transmitters to exploit all contacts with nodes belonging to a different cluster.

Once the packet arrives within the destination cluster, we can exploit well-known schemes developed for mobile network with uniform, uncorrelated mobility patterns. Indeed, notice that, under the *reshuffling* model, each cluster can be regarded as a micro-universe of nodes forming a classical mobile network in which nodes move uniformly according to the i.i.d. model. Since the throughput is bottlenecked in the previous part of the route, it turns out that, within the destination cluster, it is convenient to adopt a replication strategy, in which the packet is first broadcasted to all nodes falling within a suitable transmission range, and then one of the copies is delivered to the final destination in one more hop. This replication strategy allows to reduce the packet delivery delay without negatively impacting the overall system throughput.

Figure 3 graphically illustrates the routing scheme outlined so far. There are 4 hops and 3 intermediate relays. In the first hop, source node s sends the message to a relay node n_1 belonging to an arbitrary cluster C_r different from C_s. In the second hop, node n_1 forwards the message to a node n_2 belonging to C_d. In the third hop, the message is replicated by n_2 through a single transmission to several nodes belonging to the same cluster C_d, exploiting the intrinsic broadcast capability of the wireless channel. In the fourth hop, one of the nodes holding a copy of the message (let it be node n_3) delivers the message to the final destination.

[5]In this paper the terms packet and message are interchangeable

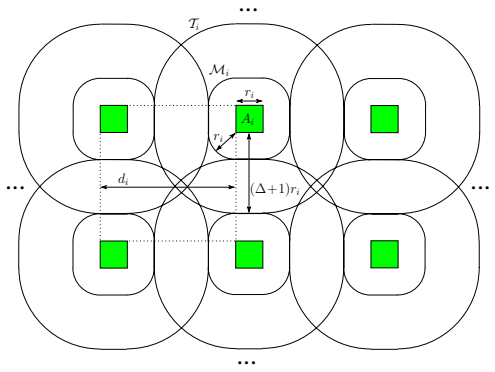

Fig. 4. Illustration of the scheduling scheme in the case of $\Delta = 2$.

B. Scheduling scheme

To implement the above described routing scheme, each node is equipped with: i) one queue storing its own generated packets (i.e., packets at hop 1); ii) $m-1$ parallel queues, one per cluster, storing packets at hop 2; iii) one queue for packets at hop 3; iv) $n/m-1$ parallel queues, one for each possible destination within its own cluster, for packets at hop 4. The service discipline is First Come First Served (FCFS) at all queues.

The scheduling scheme is in charge of selecting, at any time slot, a set of transmitter-receiver pairs which can communicate successfully according to the protocol model. Recall that the protocol model requires the adoption of the same transmission range for all communications occurring in the same slot; on the other hand, it is convenient to employ different transmission ranges for the various hops of the routing scheme. For this reason, each slot is devoted only to the transmission of packets which are at the same hop of the route. This can be equivalently done in a round-robin or in a probabilistic fashion. Following a round-robin approach, we identify every slot by a sequence number t, and in the generic slot t we allow only the transmission of packets at hop $i = |t|_4 + 1$, where $|\cdot|_m$ denotes the modulus-m operation.

One simple way to completely eliminate interference among concurrent transmissions, as required by the protocol model, is the following. Let r_i be the transmission range of packets at hop i ($i = 1, 2, 3, 4$). In any slot devoted to hop i, domain $\mathcal{O}$ is divided into squarelets $\{\mathcal{A}_i^k\}_k$ of area A_i and edge length r_i. A subset of squarelets, regularly spaced, is selected, and at most one node is allowed to transmit in each squarelet belonging to the selected subset. Figure 4 illustrates this construction for a protocol model having $\Delta = 2$. Shaded squarelets represent one possible subset of regularly spaced squarelets. Domain $\mathcal{M}_i$ around one of the squarelet denotes the maximum-size region where we can find a receiver for an arbitrary transmitter falling in the squarelet. Domain $\mathcal{T}_i$ denotes instead the region where we cannot have any other receiver belonging to a different communication pair. By spacing the selected squarelets with step $d_i = (\Delta + 2)r_i$ we can assure that one transmitter per squarelet can be enabled to transmit without generating any conflict, irrespective of the locations of transmitters within

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

their squarelets.

C. Performance analysis

To evaluate the performance of our scheme, we proceed in three steps. In Section IV-C1 we compute the maximum theoretical throughput of the system. At this stage, we assume that all queues are constantly backlogged with packets, and we compute the maximum saturation throughput achieved by inter-cluster communications, i.e., the aggregate service rate of all queues storing packets to be transmitted to nodes in different clusters. This quantity is simpler to analyze, because it requires only geometric considerations.

In Section IV-C2 we show that nodes' queues can be loaded in such a way that the actual system throughput, taking into account also traffic and queuing effects, is in order sense the same as the saturation throughput.

Then, having set the parameters of our scheme so as to maximize the system throughput, we compute in Section IV-C3 the resulting end-to-end delivery delay, showing that we cannot achieve any superior performance in terms of delay (as a secondary goal), given that we take throughput maximization as the primary goal.

1) Saturation throughput analysis: To evaluate the maximum throughput achievable in the system, we first consider the amount of data that can be simultaneously transferred in one slot between nodes belonging to different clusters. We emphasize that we are considering here arbitrary pairs of communicating nodes, with the only constraint that for each pair the transmitter and the receiver belong to different clusters.

We have the following result:

Theorem 1: In the *cluster sparse* regime, the optimal transmission range to be used for inter-cluster communications is $\Theta(R\sqrt{m/n})$. With this optimal choice, the amount of data that can be transferred in one slot among nodes belonging to different clusters is $\Theta(mR^2)$.

Proof: The proof is reported in Appendix A. ■

As immediate consequence of Theorem 1, we obtain:

Corollary 1: No scheduling-routing scheme can achieve a system throughput larger (in order sense) than mR^2.

Proof: Notice that we can neglect flows established between nodes belonging to the same cluster. Hence we can assume that all flows require at least one inter-cluster communication. A simple upper bound to the maximum achievable throughput in the system is to assume that sources have to just perform a single inter-cluster communication to an arbitrary node belonging to a different cluster, in order to deliver their packets to the destination. Let r^* be the transmission range used to perform this single hop (at any time slot). For any value of r^*, the scheduling scheme illustrated in Figure 4 is optimal, since it allows to pack (in order sense) the maximum possible number of concurrent transmissions over the network area. Then, according to Theorem 1, the best choice is to set $r^* = \Theta(R\sqrt{m/n})$, which allows to achieve an aggregate throughput $\Theta(mR^2)$. ■

2) Maximum achievable throughput: We are now ready to derive our main result on the maximum throughput achievable by our scheme:

Theorem 2: The maximum sustainable throughput of our scheme is $\Lambda = \Theta(mR^2)$ by employing a transmission range $r_i = \Theta(R\sqrt{m/n})$ for $i = 1, 2$ and $r_i = \Theta(1)$ for $i = 3, 4$. The corresponding per-node throughput is $\lambda = \Theta(mR^2/n)$.

Proof: First, we observe that by adopting the optimal transmission range $r_i = \Theta(R\sqrt{m/n})$ for inter-cluster communications, in saturated conditions, we can sustain a throughput $\Theta(mR^2)$ both at the first and second hop. Indeed, at the first hop packets are sent by construction to any node belonging to an arbitrary different cluster. In the second hop, instead, nodes have packets to transmit to nodes belonging to any cluster they come in contact with (recall that each node has $m-1$ queues associated to the second hop, one for each destination cluster).

Second, we note that: 1) the network of queues modeling the system is an acyclic network of FIFO queues (this because, by construction, every packet along its path traverses queues in increasing order of associated hop). 2) by symmetry with respect to all of the nodes, our scheme uniformly distribute the traffic among all the nodes/queues, so that all queues in the network storing packets at hop i are subject to the same ingress packet arrival rate. As a consequence of 1) and 2) all queues storing packets at hop $i = 1, 2$ are jointly stable under an arrival rate that is strictly below the service rate evaluated in saturated conditions (i.e. the saturation throughput of inter-cluster communications).

Once a packet arrives in its destination cluster, we can exploit well known schemes developed for networks with i.i.d. mobility. In principle, we could get an optimal per-node intra-cluster throughput $\Theta(1)$ by employing the 2-hop scheme of Grossglauser and Tse [2], using the same transmission range $r_i = \Theta(R\sqrt{m/n})$ adopted in previous hops, which well matches the node density within a cluster. However, this is a bad choice, because by so doing the throughput would be anyway bottlenecked by the previous hops, while we would pay excessive intra-cluster delays for nothing. Therefore in the destination cluster the optimal choice is to exactly match the throughput achievable in the previous hop, trading off capacity and delay. Indeed it is possible to enlarge the transmission range within the destination cluster up to $r_i = \Theta(1)$ without affecting the overall system throughput. This allows to adopt a replication scheme according to which the packet is forwarded in the third hop to $\Theta(\frac{n}{mR^2})$ nodes (all nodes falling within transmission range $r_3 = \Theta(1)$ of the sender) with a single broadcast transmission. Then the first node holding a copy of the message that arrives within transmission range $r_4 = \Theta(1)$ of the destination, eventually delivers the packet in hop 4. By so doing, the per-node throughput achievable at hop $i = 3, 4$ is reduced to $\Theta(\frac{mR^2}{n})$, for effect of the reduced spatial reuse, however the delay performance of the scheme is greatly improved thanks to replication (as better explained in the following delay analysis). ■

Combining Corollary 1 and Theorem 2, it immediately descends:

Corollary 2: Our proposed scheduling-routing scheme is in order sense throughput-optimal.

3) Delay Analysis: Turning our attention to the delay performance of our scheme, we focus on a particular packet belonging to a generic flow $s \rightarrow d$ (we can assume that s

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

and d belong to different clusters) and evaluate the different components of its end-to-end delivery delay, denoted by D. Let D_i be the total delay experienced by the packet at hop i. We have $D = \sum_{i=1}^{4} D_i$.

Our first step is to compute the average service time (i.e., the access delay) of the queues associated with the four hops done by the packet. Let D_i^a be the average service time of hop i. As shown in Appendix B, we have

$$\begin{array}{lll} D_1^a = \Theta\left(\frac{n}{mR^2}\right) & ; & D_2^a = \Theta\left(\frac{n}{R^2}\right) \\ D_3^a = \Theta\left(1\right) & ; & D_4^a = \Theta\left(\frac{mR^4}{n}\right) \end{array} \quad (1)$$

We observe that, at each queue, the total delay would be equal to the access delay in the absence of any contention with other packets in the network (both in the same queue and in the queues of other nodes competing for the wireless medium access). Similarly to previous work, we can shown that, at any hop, contention with other packets in the network does not change the order of magnitude of the total delay with respect to the access delay. As a result,

Theorem 3: In the *cluster sparse* regime, the delay performance of our scheme for the *reshuffling* model satisfies

$$D = \Theta\left(\max\left\{\frac{n}{R^2}, \frac{mR^4}{n}\right\}\right) \quad (2)$$

Proof:

Considering the first two hops, we observe that contention among different queues (i.e., different transmitter-receiver pairs) within the same squarelet can be neglected, since by construction only a finite number of such pairs fall w.h.p. in the squarelet. Furthermore, the queuing delay at each queue is of the same order of the access delay, because the whole system is a stable acyclic network of FIFO queues. For the analysis of the third and fourth hops, instead, we can apply previous results obtained for networks in which node movements are i.i.d [6] (recall once more that relative movements of the nodes within a cluster are i.i.d.) and claim that $D_3 = \Theta(D_3^a)$ and $D_4 = \Theta(D_4^a)$ under the condition that the injected traffic is strictly less than the saturation throughput. Since $D_i = \Theta(D_i^a)$ for all i, we have $D = \Theta(\sum_{i=1}^{4} D_i^a)$, and the result follows applying the expressions in (1). ■

We observe that similar arguments have been applied in [7], [9], [4], [6] in the case of uncorrelated i.i.d. mobility, showing that the end-to-end delay equals, in order sense, with the sum of the access delays whenever the traffic injected in the network is strictly less than the saturation throughput.

Moreover we can prove the following result, which shows that our scheme achieves optimal delay performance among the class of schemes maximizing the throughput:

Theorem 4: Any scheduling-routing scheme that achieves an aggregate throughput $\Lambda = \Theta(mR^2)$ necessarily induces a delay $D = \Omega\left(\max\left\{\frac{n}{R^2}, \frac{mR^4}{n}\right\}\right)$.

Proof: To achieve throughput $\Lambda = mR^2$ it is necessary to employ a transmission range $\Theta(R\sqrt{m/n})$ for inter-cluster communications, as a consequence of Theorem 1. Notice that using this transmission range it is not possible to get any delay gain (in order sense) by employing packet replication during inter-cluster communications (only $\Theta(1)$ nodes can simultaneously receive the message). Thus necessarily a delay $D = \Omega(D_2^a)$ must be paid, by message m to reach the destination cluster C_d, since the last relay along its path that does not belong to C_d, necessarily, has to come in contact with some node in C_d before it can transmit m. Within the destination cluster we can apply the general trade-off $D = \Omega(n\lambda^2)$ derived in [5], [6] for networks with i.i.d mobility. Using $q = n/m$ in place of n in this trade-off formula, and plugging in $\lambda = mR^2/n$, we obtain a delay $D = \Omega(\frac{mR^4}{n})$ due to intra-cluster communications. Combining to above two constraints on D due to inter- and intra- cluster communications, we get the assertion. ■

V. Cluster sparse regime: crystallized model

We describe our scheduling-routing strategy for the *crystallized* model by adapting the scheme previously defined for the *reshuffling* case. In particular, we will replace the 2-hop replication technique previously adopted within the destination cluster with a multi-hop communication similar to the one developed for static nodes by Gupta-Kumar [3]. Indeed, notice that in the *crystallized* model each cluster can be regarded as a micro-universe in which nodes are still (the relative positions of nodes within a cluster are fixed).

Let λ_M and D_M be the throughput and the delay achievable by the multi-hop communication phase performed within the destination cluster. Applying standard results for a random network of q static nodes [3], [7], a maximum per-node throughput $\hat{\lambda}_M = \sqrt{1/(q \log q)}$ can be sustained within the destination cluster (as long as sources and destinations are chosen irrespective of their locations in the cluster area), using a transmission range $\hat{r}_M = R\sqrt{\log q/q}$, at the expense of a delay $\hat{D}_M = \sqrt{q/\log q}$. Moreover, by increasing the transmission range it is possible to achieve capacity-delay trade-offs characterized by the law $D_M = \Theta(q\lambda_M)$ [7].

Similarly to the *reshuffling* model, the optimal design principle is to match the throughput achievable within the destination cluster with that provided by the inter-cluster communications performed in the previous hops, i.e., $\lambda_M = \lambda = mR^2/n$.

Depending on the system parameters, two cases are possible. If $\hat{\lambda}_M = \Omega(mR^2/n)$, which occurs for $\beta \leq (1-\nu)/4$, the intra-cluster multi-hop phase can sustain a throughput higher than or equal to the maximum system throughput $\lambda = mR^2/n$. In this case, by properly selecting the transmission range within the destination cluster we can obtain $\lambda_M = mR^2/n$, and from the trade-off law $D_M = \Theta(q\lambda_M)$ we get a corresponding delay $D_M = \Theta(R^2)$. Since $D_2^a = \Theta(n/R^2) = \omega(R^2)$ for the considered range of values of β, the overall end-to-end delay is dominated by the second hop, and we have $D = \Theta(n/R^2) = \Theta(m/\lambda)$.

If $(1-\nu)/4 < \beta < (1-\nu)/2$ (recall that we are in the *cluster sparse* regime, in which $\beta < (1-\nu)/2$), we have $\hat{\lambda}_M = o(mR^2/n)$, i.e., the last multi-hop phase, as described so far, cannot sustain the maximum throughput $\lambda = mR^2/n$ achievable by previous hops. In this case the optimal scheme is a bit trickier, as it requires to modify also the forwarding strategy of the second hop. Indeed, notice that we can increase

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

throughput λ_M beyond $\hat{\lambda}_M$ by reducing the number of multiple hops to be performed within the destination cluster, in such a way that the resulting intra-cluster throughput perfectly matches the throughput achievable in previous hops. To obtain this, we need to modify the forwarding rule of the second hop, forcing the relay node n_1 to send messages destined to d only to those nodes $n_2 \in C_d$ which fall within a proper distance $R_M = o(R)$ from d. By so doing, we reduce the length of the routes traversed by packets within the destination cluster, increasing the throughput λ_M that can be sustained by the multi-hop phase[6]. In particular, to achieve $\lambda_M = \lambda = mR^2/n$, we need to select $R_M = \Theta\left(\sqrt{\frac{q}{R^2 \log q}}\right)$. We observe that this change in the forwarding rule of the second hop requires to modify also the internal architecture of the nodes, providing each node with $n - m$ different FIFO queues, one for each destination belonging to a different cluster, in which to store packets at hop 2. Indeed, in this way we can still guarantee (in saturated traffic conditions) that, in slots devoted to hop 2, whenever a node $n_1 \in C_a$ comes in proximity of a node $n_2 \in C_b$, with $C_a \neq C_b$, it can always find a packet at the head of one queue devoted to hop 2, whose corresponding destination d lies within C_b at a distance not greater than R_M from n_2. In this way nodes can exploit all contacts with other nodes belonging to a different cluster, hence no throughput reduction occurs at hop 2 due to the modified forwarding rule.

Turning our attention to the delay of this modified scheme, the access delay of the second hop is increased to $D_2^a = \Theta(n/R_M^2)$, because a tagged packet can be forwarded only to those nodes within the destination cluster C_d which lie in a circle of radius R_M centered at the destination[7]. The delay component due to the multi-hop phase is instead equal to the number of hops $R_M/\hat{r}_M = \Theta\big(q/(R^2 \log q)\big)$. In the considered range of values for β, the end-to-end delay is always dominated by the second hop, hence $D = \Theta(mR^2 \log n) = \Theta(n\lambda \log n)$.

At last, we would like to emphasize that using similar arguments as for the *reshuffling* model, it can be proved that no scheme can achieve better delay performance in order sense, while guaranteeing the optimal throughput $\lambda = mR^2/n$.

VI. Cluster dense regime

In the *cluster dense* regime, which occurs when $\nu + 2\beta > 1$, clusters are highly overlapped at any point of the network area. Indeed, applying standard results borrowed from the theory of random geometric graphs [24], it can be shown that every point of the network area is w.h.p. covered by a number of clusters $\Theta(mR^2/n) = \Theta(n^{\nu+2\beta-1})$. This implies that nodes are almost uniformly distributed over the network domain, hence the typical distance at which one node finds the node closest to it is $\Theta(1)$. Such closest node, however, belongs w.h.p. to a different cluster. To see this, note that the density of nodes within a cluster is $\frac{q}{\pi R^2} = o(1)$, resulting into a typical distance $\omega(1)$ between nodes belonging to the same cluster.

[6]To sustain a throughput $mR^2/n = \omega(1/\log n)$, it is necessary that relay n_1 delivers the packet directly to the final destination d.

[7]The access delay D_2^a is $\Theta(n/\hat{r}_M^2)$ when packets have to be delivered directly to the final destination at hop 2.

This fact dramatically limits the degree of freedom that we have in the design of a scheduling-routing scheme specifically targeted at maximizing the system throughput. Indeed, any scheme requiring at some stage that packets are transferred between nodes belonging to the same cluster must adopt a transmission range $\omega(1)$ for such intra-cluster communications, resulting into a per-node throughput $\lambda = o(1)$.

On the contrary, a simple 2-hop scheme similar to the one proposed by Grossglauser-Tse [2], according to which packets are sent from source s to destination d though a single relay node that does not belong neither to C_s nor to C_d, can effectively employ a transmission range as short as $\Theta(1)$ (only inter-cluster communications are required), thus achieving a per-node throughput $\lambda = \Theta(1)$, at the expense of a delivery delay $D = \Theta(n)$. Such throughput-optimal scheme works for both the *reshuffling* and the *crystallized* model.

VII. Conclusions

Correlated nodes movements have huge impact on the throughput and delay performance of mobile ad hoc networks. In this paper we have provided a first characterization of the scaling laws of networks with correlated node mobility, devising novel scheduling-routing schemes which maximize the per-node throughput as primary goal. Being the first analysis of this kind, we have considered a simplified group mobility model, yet flexible enough to explore various degrees of correlation in the nodes mobility process. Our study reveals the existence of a wide range of correlated node movements which can lead to significant better performance than that achievable under independent nodes movements.

References

[1] Delay Tolerant Network Research Group, www.dtnrg.org
[2] M. Grossglauser, D.N.C. Tse, "Mobility increases the capacity of ad hoc wireless networks", *IEEE/ACM Trans. on Networking*, vol. 10, no. 2, August 2002.
[3] P. Gupta, P.R. Kumar, "The capacity of wireless networks", *IEEE Trans. on Information Theory*, 46(2), pp. 388–404, 2000.
[4] M. J. Neely, E. Modiano "Capacity and delay trade offs for ad-hoc mobile networks" *IEEE Trans. on Inf. Theory*, 51(6), pp. 1917–1937, 2005.
[5] S. Toumpis, A. J. Goldsmith, "Large Wireless Networks under Fading, Mobility, and Delay Constraints," *in Proc. INFOCOM '04*.
[6] X. Lin, N. B. Shroff, "The Fundamental Capacity-Delay Tradeoff in Large Mobile Ad Hoc Networks," *in Proc. MedHoc'04*.
[7] A. El Gamal, J. Mammen, B. Prabhakar, and D. Shah, "Throughput-Delay Trade-off in Wireless Networks", *in Proc. IEEE INFOCOM '04*.
[8] N. Bansal, Z. Liu, "Capacity, Delay and Mobility in Wireless Ad-Hoc Networks," *in Proc. IEEE INFOCOM '03*.
[9] G. Sharma, R. R. Mazumdar and N. B. Shroff, "Delay and Capacity Trade-offs in Mobile Ad Hoc Networks: A Global Perspective" *in Proc. IEEE INFOCOM '06*.
[10] J. D. Herdtner, E. K. P. Chong, "Throughput-storage tradeoff in ad hoc networks", *in Proc. INFOCOM '05*.
[11] M. Garetto, P. Giaccone, E. Leonardi, "Capacity Scaling in Delay Tolerant Networks with Heterogeneous Mobile Nodes" *in Proc. ACM MobiHoc '07*.
[12] J. Mammen, D. Shah, "Throughput and Delay in Random Wireless Networks With Restricted Mobility," *IEEE Trans. on Inf. Theory*, 53(3), pp. 1108-1116, 2007.
[13] M. Garetto, P. Giaccone, E. Leonardi, "Capacity Scaling of Sparse Mobile Ad Hoc Networks", *in Proc. INFOCOM '08*.
[14] J. Yoon, B. D. Noble, M. Liu, M. Kim, "Building realistic mobility models from coarse-grained traces", *in Proc. MobiSys '06*.
[15] V. Naumov, R. Baumann, T. Gross, "An evaluation of inter-vehicle ad hoc networks based on realistic vehicular traces," *in Proc. MobiHoc '06*.

[16] M. Musolesi, C. Mascolo, "Designing mobility models based on social network theory," *Mob. Comput. Commun. Rev.*, 11(3), pp. 59–70, 2007.
[17] P. Hui, J. Crowcroft, "Human mobility models and opportunistic communications system design," *Phil. Trans. R. Soc. A*, no. 366, pp. 2005-2016, 2008.
[18] X. Hong, M. Gerla, G. Pei, and C. Chiang. "A group mobility model for ad hoc wireless networks", *in Proc. ACM MSWiM '99*.
[19] F. Bai, N. Sadagopan, A. Helmy, "The IMPORTANT framework for analyzing the Impact of Mobility on Performance Of RouTing protocols for Adhoc NeTworks", *Ad Hoc Networks*, vol. 1, pp. 383–403, 2003.
[20] G. Pei, M. Gerla, X. Hong, "LANMAR: landmark routing for large scale wireless ad hoc networks with group mobility", *in Proc. MobiHoc '00*.
[21] M. Thomas, S. Phand, A. Gupta, "Using group structures for efficient routing in delay tolerant networks", *Ad Hoc Networks*, 7(2), pp. 344–362, 2009.
[22] A. Lindgren, A. Doria, O. Schelén, "Probabilistic Routing in Intermittently Connected Networks, *LNCS*, Vol. 3126, Springer, 2004.
[23] K. Blakely, B. Lowekamp, "A structured group mobility model for the simulation of mobile ad hoc networks", *in Proc. MobiWac '04*.
[24] M. Penrose, Random Geometric Graphs, Oxford University Press, 2003.

APPENDIX A
PROOF OF THEOREM 1

Let r_i be the transmission range used in a generic slot i devoted to inter-cluster communications among arbitrary nodes belonging to different clusters. We first observe that, according to the scheduling scheme illustrated in Figure 4, at most one communication can be enabled in each square of area d_i^2. Hence we can express the average number $\mathbb{E}[N_i]$ of packets that can be transmitted over the entire network during a slot devoted to inter-cluster communications as

$$\mathbb{E}[N_i] = \frac{n}{d_i^2}\mathbb{P}(\text{active squarelet} \mid i) \tag{3}$$

where $\mathbb{P}(\text{active squarelet} \mid i)$ is the probability that for a generic squarelet $\mathcal{A}_i^k$ and $\frac{n}{d_i^2}$ is the number of squarelets , we can find: i) a transmitting node a residing in the considered squarelet; ii) a receiving node b at distance at most r_i from a (given the location of a), and belonging to a different cluster than the one of a. Let $\mathbb{P}(a|i)$ and $\mathbb{P}(b|a,i)$ denote the occurrence probability of the two events above, respectively. Since the positions of nodes belonging to different clusters are independent, we have $\mathbb{P}(\text{active squarelet} \mid i) = \mathbb{P}(a|i)\mathbb{P}(b|a,i)$.

To evaluate $\mathbb{P}(a|i)$, we distinguish two cases. If $r_i = \Omega(R)$, a transmitting node is surely found in $\mathcal{A}_i^k$ provided that at least one cluster centre falls within it. Hence

$$\mathbb{P}(a|i) = 1 - \left(1 - \frac{r_i^2}{n}\right)^m \quad \text{for } r_i = \Omega(R) \tag{4}$$

The above expression can be approximated as $\mathbb{P}(a|i) \sim mr_i^2/n$ when $r_i^2 = O(n/m)$. Otherwise, for $r_i^2 = \omega(n/m)$, $\mathbb{P}(a|i)$ saturates to 1.

If $r_i = o(R)$, probability $\mathbb{P}(a|i)$ can be approximated by the joint occurrence of the following two events: i) $\mathcal{A}_i^k$ is entirely covered by one cluster[8]; ii) given the occurrence of event i), at least one node belonging to the covering cluster is found in $\mathcal{A}_i^k$. Condition i) above occurs when at least one cluster centre falls within a disk of radius $R - r_i/\sqrt{2}$ around the squarelet (see disk Q_a in Figure 5). While, condition ii) above occurs with probability $1 - \left(1 - \frac{r_i^2}{R^2}\right)^q$. It follows,

$$\mathbb{P}(a|i) = \left[1 - \left(1 - \frac{(R - r_i/\sqrt{2})^2}{n}\right)^m\right] \cdot \left[1 - \left(1 - \frac{r_i^2}{R^2}\right)^q\right] \quad \text{for } r_i = o(R) \tag{5}$$

The expression above can be approximated as $\mathbb{P}(a|i) \sim r_i^2$ when $qr_i^2 = O(R^2)$. Otherwise for $qr_i^2 = \omega(R^2)$ we have $\mathbb{P}(a|i) \sim mR^2/n$.

We observe that in all cases $\mathbb{P}(a|i)$ increases linearly with r_i^2 until it reaches a saturation value.

To evaluate $\mathbb{P}(b|a,i)$ in the *cluster sparse* regime, we again distinguish two cases. If $r_i = \Omega(R)$ a candidate receiver is surely found within distance r_i from a if at least one cluster center (different from the one of a) falls within a disk of radius r_i centered at a. Hence,

$$\mathbb{P}(b|a,i) = 1 - \left(1 - \frac{r_i^2}{n}\right)^{m-1} \quad \text{for } r_i = \Omega(R) \tag{6}$$

The above expression can be approximated as $\mathbb{P}(b|a,i) \sim mr_i^2/n$ when $r_i^2 = O(n/m)$. Otherwise, for $r_i^2 = \omega(n/m)$, $\mathbb{P}(b|a,i)$ saturates to 1.

If $r_i = o(R)$, probability $\mathbb{P}(b|a,i)$ can be approximated[9] by the joint occurrence of the following two events: i) the disk of radius r_i centered at a is entirely covered by a cluster different from the one of a; ii) given the occurrence of event i), at least one node belonging to the covering cluster is found in $\mathcal{A}_i^k$. Condition i) above occurs when at least one out of $m-1$ cluster centres falls within a disk of radius $R - r_i$ centered at a (see disk Q_b in Figure 6). On the other hand, condition ii)

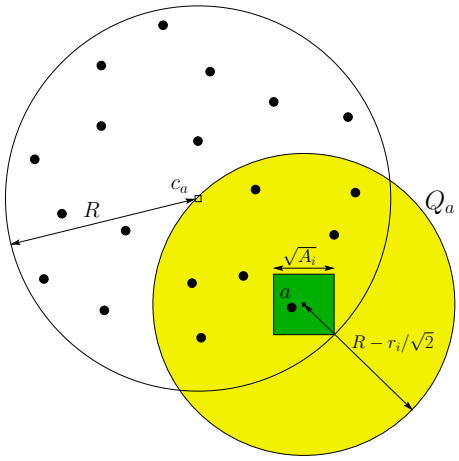

Fig. 5. The shaded disk denotes the region where a cluster center must fall so that the selected squarelet is completely covered by the cluster containing transmitting node a.

[8] a transmitter could be found in $\mathcal{A}_i^k$ even if the squarelet were partially covered by a cluster. This approximation does not affect the results, in order sense, as one can show by considering $\mathcal{A}_i^k$ entirely covered by a cluster even if just a corner of it is touched by the cluster.

[9] The approximation does not affect the result, in order sense, for reasons analogous to our approximation of $\mathbb{P}(a|i)$.

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings
This paper was presented as part of the main Technical Program at IEEE INFOCOM 2010.

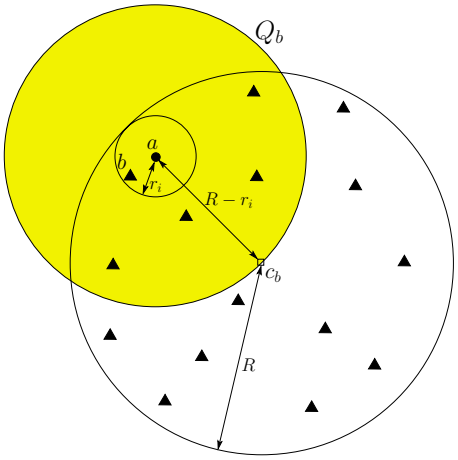

Fig. 6. The shaded disk denotes the region where a cluster center must fall so that the disk of radius r_i around transmitter a is completely covered by the cluster containing receiver b.

above occurs with probability $1 - \left(1 - \frac{r_i^2}{R^2}\right)^q$. We obtain

$$\mathbb{P}(b|a,i) = \left[1 - \left(1 - \frac{(R-r_i)^2}{n}\right)^{m-1}\right] \cdot \left[1 - \left(1 - \frac{r_i^2}{R^2}\right)^q\right] \quad \text{for } r_i = o(R) \quad (7)$$

The expression above can be approximated as $\mathbb{P}(b|a,i) \sim r_i^2$ when $qr_i^2 = O(R^2)$. Otherwise for $qr_i^2 = \omega(R^2)$ we have $\mathbb{P}(b|a,i) \sim mR^2/n$.

We observe again that $\mathbb{P}(b|a,i)$ increases linearly with r_i^2 until it reaches a saturation value.

Given that $d_i^2 = \Theta(r_i^2)$, putting things together we obtain:

$$\mathbb{E}[N_i] = \begin{cases} \frac{(mr_i)^2}{n} & r_i = \Omega(R)\ ,\ r_i = O(\sqrt{n/m}) \\ \frac{n}{r_i^2} & r_i = \Omega(R)\ ,\ r_i = \omega(\sqrt{n/m}) \\ nr_i^2 & r_i = o(R)\ ,\ r_i = O(R\sqrt{m/n}) \\ \frac{m^2R^4}{nr_i^2} & r_i = o(R)\ ,\ r_i = \omega(R\sqrt{m/n}) \end{cases} \quad (8)$$

From (8) we conclude that if $r_i = \Omega(R)$ we can obtain at most $\mathbb{E}[N_i] = m$ by choosing $r_i = \sqrt{n/m}$. This corresponds to using a transmission range equal to the typical distance between cluster centers. Instead, if $r_i = o(R)$ it is possible to achieve at most $\mathbb{E}[N_i] = mR^2$ by selecting $r_i = R\sqrt{m/n}$. This corresponds to using a transmission range which is strictly related to the density of nodes within clusters, which is equal to $n/(mR^2)$. In particular, with this choice r_i is equal to the typical distance between nodes belonging to the same cluster.

APPENDIX B
COMPUTATION OF THE ACCESS DELAYS

Since we are interested to an order sense evaluation of the access delay of each hop, we can ignore all factors whose effect on the access delay can be bounded by a multiplicative constant, such as: i) the fact that only one slot out of four is devoted to transmission of packets at a given hop; ii) only a subset of squarelets can be activated in a given slot.

In the first hop, the tagged packet has to wait until the source node s gets in contact with a node n_1 belonging to an arbitrary different cluster C_r (Figure 3). In a slot devoted to hop 1, the probability $\mathbb{P}(n_1|s,1)$ that a generic node n_1 belonging to a cluster C_r different from C_s gets in contact with s (i.e., lies at distance at most r_1 from s) is analogous to quantity $\mathbb{P}(b|a,i)$ in (7) (see Figure 6). Since we use $r_1 = R\sqrt{m/n}$, we have $\mathbb{P}(n_1|s,1) \sim R^2m/n$.

The packet access delay at s, expressed in number of slots, follows a geometric distribution Geom($\mathbb{P}(n_1|s,1)$), since the positions of all nodes regenerate from slot to slot. Hence the average access delay at the fist hop is:

$$D_1^a = \Theta\left(\frac{n}{mR^2}\right)$$

The access delay in the second hop is similar to the one of the first hop, however in this case n_1 can transmit the tagged packet only when it gets in contact with a node n_2 belonging to the *specific* cluster containing the destination. In a slot devoted to hop 2, the probability $\mathbb{P}(n_2|n_1,2)$ that n_1 gets in contact with a generic node n_2 belonging to cluster C_d can be computed as:

$$\mathbb{P}(n_2|n_1,2) = \frac{(R-r_2)^2}{n}\left[1 - \left(1 - \frac{r_2^2}{R^2}\right)^q\right]$$

Since we use $r_2 = R\sqrt{m/n}$, we have $\mathbb{P}(n_2|n_1,2) \sim R^2/n$. Again, the packet access delay at n_1, expressed in number of slots, follows a geometric distribution Geom($\mathbb{P}(n_2|n_1,2)$), thus:

$$D_2^a = \Theta\left(\frac{n}{R^2}\right)$$

Note that D_2^a always dominate D_1^a.

To compute D_3^a and D_4^a we can apply standard results [7], [9], [6], [5] obtained for the i.i.d mobility model, since relative movements of nodes within each cluster area are i.i.d.

In particular, we have $D_3^a = \Theta(1)$, since n_2 can broadcast the packet in any slot devoted to the third hop without any other requirement.

After the third hop, a number $\Theta(\frac{n}{mR^2})$ of nodes within the destination cluster hold a copy of the tagged packet, hence D_4^a corresponds to the average time that it takes before the first one of these nodes arrive at distance $r_4 = \Theta(1)$ from the destination. In a slot devoted to hop 4, the probability $\mathbb{P}(n_3|d,4)$ that at least one node holding a copy of the tagged packet falls within transmission range r_4 from d is given by

$$\mathbb{P}(n_3|d,4) = 1 - \left(1 - \frac{r_4^2}{R^2}\right)^{\frac{n}{mR^2}} = \Theta\left(\frac{n}{mR^4}\right)$$

Since the access delay of the fourth hop follows a geometric distribution Geom($\mathbb{P}(n_3|d,4)$), we have:

$$D_4^a = \Theta\left(\frac{mR^4}{n}\right)$$