Capacity Region, Minimum Energy and Delay for a Mobile Ad-Hoc Network

Rahul Urgaonkar, Michael J. Neely
Department of Electrical Engineering-Systems
University of Southern California, Los Angeles, CA 90089
urgaonka, mjneely@usc.edu

***Abstract*— We investigate two quantities of fundamental interest in a mobile ad-hoc network: the capacity region and the minimum energy function of the network. The capacity region is defined as the closure of the set of all input rates that the network can stably support. The minimum energy function establishes a lower bound on the amount of energy required to support a given set of input rates. We consider a specific model of the mobile ad-hoc network that enables us to exactly compute these quantities. Further, we propose schemes that offer performance guarantees that are arbitrarily close to these bounds at the cost of an increased delay. The exact nature of the associated delay tradeoff when performance is pushed towards the minimum energy bound is another fundamental characteristic of the network that is discussed in this work.**

I. Introduction

Two quantities that characterize the performance limits of a mobile ad-hoc network are the capacity region and the minimum energy function of the network. The capacity region is defined as the closure of the set of all possible input rates that the network can stably support considering all possible scheduling and routing algorithms that conform to the given network structure. The minimum energy function can be defined as the minimum time average energy (summed over all users) required to stably support a given set of input rates. It is, therefore, not possible for any scheme to support an input rate at an energy cost lower than the minimum energy function value. In this work, we compute these quantities exactly for a specific model of a mobile ad-hoc network. We then describe schemes whose performance can be pushed arbitrarily close to these bounds at the cost of an increased delay.

Asymptotic bounds on the capacity of static wireless networks and mobile networks are developed by [3], [4]. Capacity-delay tradeoffs in such networks are considered in [1], [2], [6], [7], [8]. However, little work has been done in computing the *exact* capacity and energy expressions for these networks. Exceptions include a closed form expression for the capacity of a fixed grid network in [5], and an expression for the capacity of a mobile ad-hoc network in [1], [2]. Specifically, the work in [1], [2] uses a cell-partitioned structure and assumes only same-cell transmissions are possible and that only a single packet can be sent on each transmission. In this work, we extend this model to more general scenarios allowing adjacent cell communication and different rate-power combinations. Specifically, the contributions of this paper are threefold:

1) We extend the simplified cell-partitioned model of [1], [2] to treat inter-cell communication. We establish exact capacity expressions for general user mobility processes (possibly non-i.i.d. and non-uniform), assuming only a well-defined steady state location distribution for the users. Our analysis shows that the throughput optimal solution for this extended model is also necessarily a 2-hop relay algorithm. Further, our analysis illuminates the optimal decision strategies and precisely defines the throughput optimal control law for choosing between *same* cell and *adjacent* cell transmission. We then use this insight to design a 2-hop relay algorithm that can stabilize the network for all input rates within the capacity region. We also compute an upper bound on the delay of this algorithm. Our analytical technique can be extended to systems with additional constraints as well. (Section III)
2) We compute the exact expression for the minimum energy required to stabilize this network, for all input rates within capacity. Our result demonstrates a piecewise linear structure for the minimum energy function. (Section IV)
3) We present a *greedy* algorithm whose energy cost can be pushed arbitrarily close to the minimum energy at the cost of an increased delay. We then discuss the energy-delay tradeoff involved and note that the piecewise linear special structure of the minimum energy function can be exploited to design a multi-hop policy that achieves an energy-delay tradeoff superior to that given in [10], and even outperform the square-root tradeoff law given in [11]. (Section V)

Our network model and assumptions are discussed in the next section.

II. Model and Assumptions

1) Network Model: We use the following cell partitioned network model: The network is partitioned into C non-overlapping cells, not necessarily of the same size/shape (see Fig. 1). There are N users independently roaming from cell to cell over the network. Note that there could be "gaps" in the cell structure due to infeasible geographic locations. We assume that the gaps do not

This material is based on work supported in part by the National Science Foundation under grant OCE 0520324.

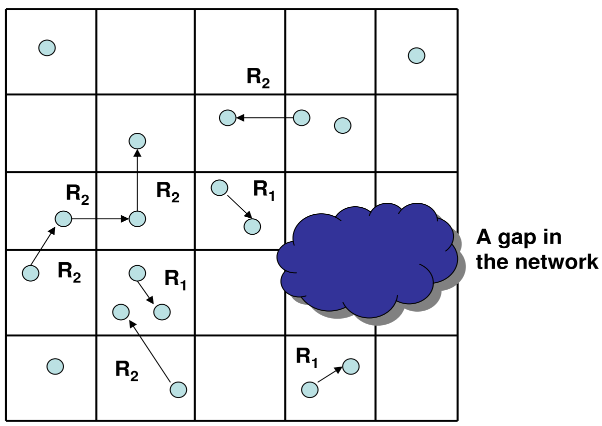

Fig. 1. An illustration of the cell-partitioned network with same and adjacent cell communication and gaps in the structure

partition the network, so that it is possible for a single user to visit all cells. We assume C is the number of valid cells, not including such gaps. We can then define the user density $d = \frac{N}{C}$ users/cell.

2) Mobility Model: Time is slotted so that each user remains in its current cell for a timeslot and potentially moves to a new cell at the end of the slot. We assume that each user moves *independently* of the other users according to a *general* mobility process. We assume only that the mobility process has a well-defined steady state location distribution π_c over the cells $c \in \{1, 2, ..., C\}$. This distribution could be non-uniform over the cells. Thus, for example, our analysis can be used to compute the exact capacity and minimum energy for a network in which users are performing a Markovian Random Walk over the cells such that there exists a well-defined steady state location distribution for all users. We assume the steady state cell location distribution π_c is the same for all users.

   We assume that the mobility model is ergodic and that time average location probaiblities converge to their steady state location probabilities π_c. We further assume that the mobility model has the following "renewable" property: Given any $\delta > 0$, there exists a finite integer K such that the expected time average location probability taken over any interval of size K is within δ of its long term time average. That is, for any time t:

$$\frac{1}{K}\sum_{\tau=0}^{K-1} \text{Pr[user in cell c at time } t+\tau | H(t)] - \pi_c \leq \delta$$

   regardless of past history up to time t, where $H(t)$ is the history up to time t.

3) Traffic Model: Packets are assumed to arrive at a user i according to some arrival process $A_i(t)$ which indicates the number of packet arrivals in timeslot t for user i. λ_i represents the rate of this process, i.e.

$$\lim_{t\to\infty} \frac{\sum_{\tau=0}^{\tau=t-1} A_i(\tau)}{t} = \lambda_i \qquad \text{w.p. } 1$$

   We assume that the $A_i(t)$ variables are i.i.d and $\mathbb{E}\{A_i(t)\} = \lambda_i$. We also assume that packets generated by user i are destined for a unique user j and vice versa. For simplicity, we assume N is even with the following one-to-one pairing between users: $1 \leftrightarrow 2, 3 \leftrightarrow 4, ..., (N-1) \leftrightarrow N$, i.e., packets generated by user 1 are destined for user 2 and those generated by user 2 are destined for user 1 and so on.

4) Communication Model: We assume that two users can communicate *only if* they are in the same cell or in an adjacent cell. Further, if the communication takes place in the same cell, R_1 packets can be transmitted from the sender to the receiver if the sender uses full power. If the receiver is in an adjacent cell, R_2 packets can be transmitted with full power. We assume $R_1 \geq R_2$. Power allocation is restricted to the set $\{0, 1\}$, i.e., each user either uses zero power or full power (we use normalized value). We allow at most one transmitter in a cell at any given timeslot, though it may have multiple receivers (due to possible adjacent cell communication). Further, a user may potentially transmit and receive simultaneously. This model is conceivable if the users in neighboring cells use orthogonal communication channels.

This model allows us to treat scheduling decisions in each cell *independently* of all other cells. It is possible to include additional constraints in the communication model, e.g., only 1 receiver per cell, no simultaneous transmission and reception etc. We note that this model can easily be extended to the scenarios where users can choose from a finite set of power allocations. Similarly, our model can also be extended to take into account channel states on *each* communication link. We discuss these extensions at the end of Sec.III.

In this work, we restrict our attention to network control algorithms that operate according to the given network structure described above. A general algorithm within this class will make scheduling decisions about what packet to transmit, when, and to whom. For example, it may decide to transmit to a user in an adjacent cell rather than to some user in the same cell, even though the transmission rate is smaller. However, we assume that the packets themselves are kept intact and are not combined or network coded.

III. Capacity, Delay and a 2-hop Relay Algorithm

In this section, we compute the exact capacity of the network model previously described. We assume that all users receive packets at the same rate (i.e. $\lambda_i = \lambda$ for all i)[1]. The capacity of the network is then the maximum rate λ that the network can stably support.

Theorem 1: The capacity of the network is:

$$\mu = \begin{cases} \frac{2R_1q + R_1(p-q) + 2R_2q' + R_2(p'-q')}{2d} & \text{if } R_1 \geq 2R_2 \\ \frac{2R_1q + R_1p'' + 2R_2q'' + R_2(p'-q')}{2d} & \text{if } 2R_2 > R_1 \geq R_2 \end{cases}$$

[1]The scenario involving heterogeneous rates can be treated similarly.

where
$q = \frac{1}{C}\sum_{c=1}^{C}$Pr[finding a S-D pair in a cell c]
$p = \frac{1}{C}\sum_{c=1}^{C}$Pr[finding at least 2 users in a cell c]
$q' = \frac{1}{C}\sum_{c=1}^{C}$Pr[finding exactly 1 user in a cell c and its destination in an adjacent cell]
$p' = \frac{1}{C}\sum_{c=1}^{C}$Pr[finding exactly 1 user in a cell c and at least 1 user in an adjacent cell]
$q'' = \frac{1}{C}\sum_{c=1}^{C}$Pr[finding no S-D pair in a cell c but at least 1 S-D pair with an adjacent cell]
$p'' = \frac{1}{C}\sum_{c=1}^{C}$Pr[finding no S-D pair in a cell c and any adjacent cell but at least 2 users in the cell c]

Note that the value of μ does not vanish with the number of users N as long as the network density is $O(1)$, i.e., the number of cells C is scaled appropriately as N grows. This is in agreement with the $O(1)$ per user throughput for a mobile ad-hoc network, first observed in [4].

The probabilities in the summations above are the steady state probabilities associated with the steady state cell location distributions of the users. Thus, using the steady state cell location distribution π_c, we can exactly compute these probabilities for our network. These are given by:

$$q = \frac{1}{C}\sum_{c=1}^{C}\Big(1-(1-\pi_c^2)^{N/2}\Big)$$

$$p = \frac{1}{C}\sum_{c=1}^{C}\Big(1-(1-\pi_c)^N - N\pi_c(1-\pi_c)^{N-1}\Big)$$

$$q' = \frac{1}{C}\sum_{c=1}^{C}\Big(\Pi_{adj}(c)N\pi_c(1-\pi_c)^{N-1}\Big)$$

$$p' = \frac{1}{C}\sum_{c=1}^{C}\Big(1-(1-\Pi_{adj}(c))^{N-1}\Big)N\pi_c(1-\pi_c)^{N-1}$$

$$q'' = \frac{1}{C}\sum_{c=1}^{C}\sum_{i=1}^{N/2}2^i\binom{N/2}{i}\pi_c^i(1-\pi_c)^{N-i}\Big(1-(1-\Pi_{adj}(c))^i\Big)$$

$$p'' = \frac{1}{C}\sum_{c=1}^{C}\sum_{i=2}^{N/2}2^i\binom{N/2}{i}\pi_c^i(1-\pi_c)^{N-i}(1-\Pi_{adj}(c))^i$$

Here, $\Pi_{adj}(c)$ denotes the sum of the conditional steady state probabilities of a user being in any adjacent cell of cell c given that this user is not in cell c, i.e., $\Pi_{adj}(c) = \frac{1}{1-\pi_c}\sum_{i=1}^{C}\pi_c 1_{\text{i adjacent to c}}$ where $1_{\text{i adjacent to c}}$ is an indicator function taking value 1 if cell i is adjacent to cell c and 0 otherwise.

In the expressions for q'', p'', the term $2^i\binom{N/2}{i}\pi_c^i(1-\pi_c)^{N-i}$ is the probability of finding i users in a cell such that there are no source-destination pairs. It can be obtained by noting that $2^i\frac{\binom{N/2}{i}}{\binom{N}{i}}$ is the probability of finding no source-destination pair is a cell *given that* i users are in that cell and $\binom{N}{i}\pi_c^i(1-\pi_c)^{N-i}$ is the probability of choosing i out of N users. Clearly, $i \leq \frac{N}{2}$ since there must be at least 1 source-destination pair for $i > \frac{N}{2}$.

As an example, if we consider the case when the steady state location distribution is uniform over all cells (so that $\pi_c = \frac{1}{C}$ for all c), then we have the following simpler expressions for q, p, q', p':

$$q = 1-(1-\frac{1}{C^2})^{N/2}$$

$$p = 1-(1-\frac{1}{C})^N - \frac{N}{C}(1-\frac{1}{C})^{N-1}$$

$$q' = \frac{N}{C^2}(1-\frac{1}{C})^{N-1}\sum_{c=1}^{C}\Pi_{adj}(c)$$

$$p' = \frac{N}{C^2}(1-\frac{1}{C})^{N-1}\sum_{c=1}^{C}\Big(1-(1-\Pi_{adj}(c))^{N-1}\Big)$$

Assuming each cell has at most 4 adjacent cells (to the left, right, up and down respectively), an example of a mobility process that would give a uniform steady state location distribution under our cell-partitioned network model is a random walk where every time slot, each user stays in the previous cell with some probability $1-\beta$ (such that $0 < \beta < 1$), else decides to move to an adjacent cell with probability $\beta/4$. If there is no feasible adjacent cell (e.g, if the previous cell is a corner cell and the new chosen cell doesn't exist), then the user remains in the previous cell. This random walk forms an irreducible and aperiodic Markov Chain since we assume that there are no partitions in the network and there are self-transitions with positive probability.

Thus, the network can stably support users simultaneously communicating at any rate $\lambda < \mu$. We prove the theorem in 2 parts. First, we establish the necessary condition by deriving an upper bound on the capacity of *any* stabilizing algorithm. Then, we establish sufficiency by presenting a specific scheduling strategy and showing that the average delay is bounded under that strategy.

Proof of Necessity: Let S be the set of all stabilizing scheduling policies. Consider any particular policy $s \in S$. Suppose it successfully delivers $X_{ab}^s(T)$ packets from sources to destinations involving "a" same cell transmissions and "b" adjacent cell transmissions in the interval $(0,T)$. Fix $\epsilon > 0$. For stability, there must exist arbitrarily large values of T such that the total output rate is within ϵ of total input rate. Thus:

$$\frac{\sum_{a=0}^{\infty}\sum_{b=0}^{\infty}X_{ab}^s(T)}{T} \geq N\lambda - \epsilon \tag{1}$$

Next, total number of packet transmissions in $(0,T)$ is *at least* $\sum_{a=0}^{\infty}\sum_{b=0}^{\infty}(a+b)X_{ab}^s(T)$ (because these many packets were certainly delivered). This must be less than the maximum possible packet transmission opportunities defined as $Y^s(T)$. Thus:

$$\begin{aligned}\frac{1}{T}Y^s(T) &\geq \frac{1}{T}\sum_{a=0}^{\infty}\sum_{b=0}^{\infty}(a+b)X_{ab}^s(T)\\ &\geq \frac{1}{T}\sum_{a+b<2}X_{ab}^s(T) + \frac{2}{T}\sum_{a+b\geq 2}X_{ab}^s(T)\\ &\geq \frac{1}{T}\sum_{a+b<2}X_{ab}^s(T)\\ &\quad + \frac{2}{T}\Big((N\lambda-\epsilon) - \sum_{a+b<2}X_{ab}^s(T)\Big)\text{(using (1))}\end{aligned}$$

Hence, noting that ϵ can be chosen to be arbitrarily small, we have:

$$\lambda \leq \lim_{T\to\infty} \frac{Y^s(T) + X^s_{10}(T) + X^s_{01}(T)}{2TN} \tag{2}$$

Define $Y^s_c(\tau)$ as the number of packet transmission opportunities in cell c at time τ. Then $Y^s(T) + X^s_{10}(T) + X^s_{01}(T)$ can be written as a sum over all time slots τ and all cells c as follows:

$$\begin{aligned}
&\sum_{\tau=0}^{T-1} \Big(Y^s(\tau) + X^s_{10}(\tau) + X^s_{01}(\tau)\Big)\\
&= \sum_{\tau=0}^{T-1}\sum_{c=1}^{C} \Big(Y^s_c(\tau) + X^s_{10,c}(\tau) + X^s_{01,c}(\tau)\Big)\\
&\leq \sum_{\tau=0}^{T-1}\sum_{c=1}^{C} \max_{s\in S} \Big(Y^s_c(\tau) + X^s_{10,c}(\tau) + X^s_{01,c}(\tau)\Big) \qquad (3)
\end{aligned}$$

Define $Z^s_c(\tau) = Y^s_c(\tau) + X^s_{10,c}(\tau) + X^s_{01,c}(\tau)$ and $\hat{Z}^c(\tau) = \max_{s\in S} Z^s_c(\tau)$. Also, define the following indicator decision variables for any policy s for some $\tau \in [0, T]$ and $c \in [1, C]$:

$$X^c_1 = \begin{cases} 1 & \text{if } s \text{ schedules a same cell direct} \\ & \text{transmission at rate } R_1 \text{ in cell c} \\ 0 & \text{else} \end{cases}$$

$$X^c_2 = \begin{cases} 1 & \text{if } s \text{ schedules a same cell relay} \\ & \text{transmission at rate } R_1 \text{ in cell c} \\ 0 & \text{else} \end{cases}$$

$$X^c_3 = \begin{cases} 1 & \text{if } s \text{ schedules an adjacent cell direct} \\ & \text{transmission at rate } R_2 \text{ in cell c} \\ 0 & \text{else} \end{cases}$$

$$X^c_4 = \begin{cases} 1 & \text{if } s \text{ schedules an adjacent cell relay} \\ & \text{transmission at rate } R_2 \text{ in cell c} \\ 0 & \text{else} \end{cases}$$

Thus, we can express $Z^s_c(\tau)$ as follows:

$$\begin{aligned}
Z^s_c(\tau) &= Y^s_c(\tau) + X^s_{10,c}(\tau) + X^s_{01,c}(\tau)\\
&= R_1X^c_1 + R_1X^c_2 + R_2X^c_3 + R_2X^c_4 + X^c_{10}(\tau) + X^c_{01}(\tau)\\
&= 2R_1X^c_1 + R_1X^c_2 + 2R_2X^c_3 + R_2X^c_4
\end{aligned}$$

Since only one of the decision variables is 1 and the rest are 0, the preference order for decisions to maximize $Z^s_c(\tau)$ is evident. Specifically, it would be $X^c_1 \prec X^c_2 \prec X^c_3 \prec X^c_4$ when $R_1 \geq 2R_2$ and $X^c_1 \prec X^c_3 \prec X^c_2 \prec X^c_4$ when $R_2 \leq R_1 < 2R_2$. So in each cell c, $Z^s_c(\tau)$ is maximized by choosing the decisions in this preference order, only choosing a less preferred decision when none of the more preferred decisions are possible in that cell.

Thus, (from (2) and (3)):

$$\lambda \leq \lim_{T\to\infty} \frac{1}{2TN} \sum_{\tau=0}^{T-1}\sum_{c=1}^{C} \hat{Z}^c(\tau)$$

As $\hat{Z}^c(\tau)$ can take only a finite number of values (namely $R_1, R_2, 2R_1, 2R_2, 0$) and is a function of the current state of the ergodic user location processes, the time average of $\hat{Z}^c(\tau)$ is exactly equal to its expectation with respect to the steady state user location distribution.

Thus, the capacity region can now be computed by calculating the expectation of $\hat{Z}^c$ using the steady state probabilities associated with the indicator variables and summing over all cells. These are given by q, p, q', p', q'', p'' for both cases of interest. □

Note that the above preference order clearly spells out the structure of the throughput optimal strategy. Specifically, depending on the values of R_1 and R_2, it can be used to decide between same cell relay versus adjacent cell direct transmission. We use this structure of the throughput optimal strategy to design a 2-hop relay algorithm that we present next.

Also note the factor of 2 with the decision variables corresponding to direct source-destination transmission. Intuitively, each such transmission opportunity is better than a similar opportunity between source-relay or relay-destination by a factor of 2 since the indirect transmissions need twice as many opportunities to deliver a given number of packets to the destination as compared to direct transmissions.

Proof of Sufficiency: Now we present an algorithm that makes stationary randomized scheduling decisions and show that it gives bounded delay for any rate $\lambda < \mu$, i.e., there exists $\epsilon > 0$ such that $\lambda + \epsilon \leq \mu$. We only consider the case when $R_1 \geq 2R_2$. The other case is similar and is not discussed.

2 Hop Relay Algorithm: Every timeslot, for all cells, do the following:

1) If there exists a source-destination pair in the cell, randomly choose such a pair (uniformly over all such pairs in the cell). If the source has new packets for the destination, transmit at rate R_1 (add dummy packets if less than R_1 packets present). Else remain idle.
2) If there is no source-destination pair in the cell but there are at least 2 users in the cell, randomly designate one user as the sender and another as the receiver. Then toss an *unfair* coin such that the probability of "Head" is $\frac{1-\delta}{2}$ where $\delta = \delta(\epsilon)$ and $0 < \delta < 1$. If the outcome is "Head", perform the first action below. Else, perform the second.
   a) *Send new Relay packets in same cell:* If the transmitter has new packets for its destination, transmit at rate R_1 (add dummy packets if less than R_1 packets present). Else remain idle.
   b) *Send Relay packets to their Destination in same cell:* If the transmitter has packets for the receiver, transmit at rate R_1 (add dummy packets if less than R_1 packets present). Else remain idle.
3) If there is only 1 user in the cell and its destination is present in one of the adjacent cells, transmit at rate R_2 if new packets present (add dummy packets if less than R_2 packets present). Else remain idle.
4) If there is only 1 user in the cell and its destination is not present in one of the adjacent cells but there is at least one user in an adjacent cell, randomly designate one such user as the receiver and the only user in the cell as the transmitter. Then toss an *unfair* coin such that the probability of a "Head" is $\frac{1-\delta}{2}$ where $\delta = \delta(\epsilon)$ and

$0 < \delta < 1$. If the outcome is a "Head", perform the first action below. Else, perform the second.

a) *Send new Relay packets in adjacent cell:* If the transmitter has new packets for its destination, transmit at rate R_2 (add dummy packets if less than R_2 packets present). Else remain idle.

b) *Send Relay packets to their Destination in adjacent cell:* If the transmitter has packets for the receiver, transmit at rate R_2 (add dummy packets if less than R_2 packets present). Else remain idle.

This algorithm is motivated by the proof of Theorem 1 since it follows the same preference order in making scheduling decisions. Note that this algorithm restricts the path lengths of all packets to *at most* 2 hops because any packet that has been transmitted to a relay node is restricted from being transmitted to any other node except its destination.

We make use of the following Lyapunov Drift Lemma (from [1]) to analyze the performance of this 2 Hop Relay Algorithm. Consider a network of N queues operating in slotted time, and let $\underline{U}(t) = (U_1(t), U_2(t), ..., U_N(t))$ represent a row vector of unfinished work in each of the queues for timeslots $t \in \{0, 1, 2, ...\}$. Here, the unfinished work $U_i(t)$ represents the number of backlogged packets destined for user i. Define a non-negative function $L(\underline{U}(t))$ of the unfinished work vector $\underline{U}(t)$ called a Lyapunov function. Then we have the following:

Lyapunov Drift Lemma: If there exists a positive integer K such that for all timeslots t and for all $\underline{U}(t)$, the Lyapunov function of unfinished work $L(\underline{U})$ evaluated K steps into the future satisfies:

$$\mathbb{E}\{L(\underline{U}(t+K)) - L(\underline{U}(t))|\underline{U}(t)\} \leq B - \sum_{i,c} \theta_i^{(c)} U_i^{(c)}(t)$$

for some positive constants B, $\{\theta_i^{(c)}\}$, and if $\mathbb{E}\{L(\underline{U}(t))\} < \infty$ for $t \in \{0, 1, ..., K-1\}$, then the network is stable, and:

$$\limsup_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \Big[\sum_i \theta_i \mathbb{E}\{U_i(\tau)\} \Big] \leq B$$

Proof: This Lemma can be shown by using a telescoping series argument and is proved in [1].

Theorem 2: For a cell partitioned network (with N nodes and C cells) as described above, with a capacity μ and input rates λ for each user such that $\lambda + \epsilon \leq \mu$ for some $\epsilon > 0$, the average packet delay $\overline{D}$ under the 2 Hop Relay Algorithm satisfies:

$$\overline{D} \leq \frac{O(KN)}{\epsilon}$$

where K is a finite integer (that may depend on N) related to the mixing time of the mobility process. i.e., the time required for the mobility process to reach a "near steady state", described in more detail in [1].

Proof: Let $U_i^{(c)}(t)$ represent the amount of unfinished work of type c (i.e. number of packets destined for node c) that are queued up in node i at time t. The K-step dynamics of unfinished work satisfies the following condition for all $i \neq c$:

$$U_i^{(c)}(t+K) \quad \leq \quad \max\Big(U_i^{(c)}(t) - \sum_{\tau=t}^{t+K-1} \sum_b \mu_{ib}^{(c)}(\tau), 0\Big) + \sum_{\tau=t}^{t+K-1} \sum_a \mu_{ai}^{(c)}(\tau) + \sum_{\tau=t}^{t+K-1} A_{ic}(\tau) \qquad (4)$$

where
$A_{ic}(\tau)$ = number of new type c arrivals to source node i at the beginning of time slot τ
$\mu_{ab}(\tau)$ = rate offered to type c packets in time slot τ with node a as transmitter and node b as receiver

The above condition is an inequality because the total arrivals to node i from other nodes may be less than $\sum_{\tau=t}^{t+K-1} \sum_a \mu_{ai}^{(c)}(\tau)$ if these other nodes have little or no packets to send.

Now define the Lyapunov function $L(\underline{U}) = \sum_{i\neq c}[U_i^{(c)}]^2$. Then we have the following expression for the K slot Lyapunov Drift (from [1]):

$$\begin{aligned}
&\mathbb{E}\{L(\underline{U}(t+K)) - L(\underline{U}(t))|\underline{U}(t)\} \leq K^2BN \\
&- 2K \sum_{i\neq c} U_i^{(c)}(t) \frac{1}{K} \sum_{\tau=t}^{t+K-1} \mathbb{E}\Big\{ \sum_b \mu_{ib}^{(c)}(\tau) \\
&- \sum_a \mu_{ai}^{(c)}(\tau) - A_{ic}(\tau)|\underline{U}(t)\Big\} \qquad (5)
\end{aligned}$$

where
$B = (A_{max} + \mu_{max}^{in})^2 + (\mu_{max}^{out})^2$
A_{max}^2 = upper bound on the second moment of the exogenous arrivals to any node
μ_{max}^{in} = maximum transmission rate into any node
μ_{max}^{out} = maximum transmission rate out of any node

Now note that $\exists K$ such that $\frac{1}{K}\sum_{\tau=t}^{t+K-1} \mathbb{E}\Big\{ \sum_b \mu_{ib}^{(c)}(\tau) - \sum_a \mu_{ai}^{(c)}(\tau) - A_{ic}(\tau)\Big\}$ is within $\delta/2$ of the time average values $\sum_b \overline{\mu}_{ib}^{(c)} - \sum_a \overline{\mu}_{ai}^{(c)} - \lambda_{ic}$. This follows from the assumptions on ergodicity and renewability in the mobility model (see Sec. II) by noting that the functionals of user cell locations also converge over finite intervals of size K.

The drift condition holds for any control policy that yields general $\mu_{ij}^{(c)}(t)$ rates. We next calculate the rates obtained by the 2 Hop Relay Algorithm. Since this algorithm makes transmission decisions independent of the current backlog levels, we can ignore the conditional expectation in the last summation above. We have the following 2 cases[2] :

1) Node i generates type c packets: In this case, $\mathbb{E}\{A_{ic}(\tau)\} = \lambda$ and $\mathbb{E}\{\sum_a \mu_{ai}^{(c)}(\tau)\} = 0$ (since a source node would never get back a packet that it generates). To calculate $\frac{1}{K}\sum_{\tau=t}^{t+K-1} \mathbb{E}\{\sum_b \mu_{ib}^{(c)}(\tau)\}$, we note that due to the randomized nature of the 2 Hop Relay Algorithm, packets from a given source node see the network only as a source, destination and intermediate relays and transmissions of packets from other sources are reflected

[2] The case where node i is the receiver for type c packets doesn't arise since the summation is over $i \neq c$

simply as random ON/OFF service opportunities. Therefore, the outgoing service rate for packets generated by the source is equal to the sum of the rate at which the source is scheduled to transmit directly to its destination and the rate at which it is scheduled to transmit to one of the relay users. Let these rates be r_1 and r_2 respectively. Also let the transmission rate out of the relay nodes be r_3. Then the total rate of transmission over the network is $N(r_1 + r_2 + r_3)$.
Using the probability of choosing source-relay and relay-destination transmissions, we have

$$r_2 = \frac{1-\delta}{1+\delta} r_3 \tag{6}$$

In the 2 Hop Relay Algorithm, a direct source-to-destination transmission is scheduled whenever there is a source-destination pair is the same cell or there is only 1 node is a cell and its destination is in an adjacent cell. Thus, using the result from the proof of Theorem 1, we have:

$$C(R_1 q + R_2 q') = N r_1 \tag{7}$$

Similarly, the sum total transmissions in the network can be expressed in terms of the quantities p and p' as follows:

$$C(R_1 p + R_2 p') = N(r_1 + r_2 + r_3) \tag{8}$$

Using these, we have:

$$r_1 + r_2 = \frac{(1+\delta)(R_1 q + R_2 q') + (1-\delta)(R_1 p + R_2 p')}{2d}$$

By definition, K is the number of slots required for the mobility process to reach a near steady state. Since the total outgoing rate is only a function of the steady state location distribution, $\frac{1}{K}\sum_{\tau=t}^{t+K-1} \mathbb{E}\{\sum_b \mu_{ib}^{(c)}(\tau)\}$ is within $\delta/2$ of $r_1 + r_2$ as obtained above.
Choosing $\delta = \frac{d\epsilon}{R_1 p + R_2 p' - R_1 q - R_2 q' - d}$, the above expression is at least $\mu - \epsilon/2$ where μ is the capacity of the network. Thus, the last summation in (5) for this case is at least $\mu - \epsilon/2 - \lambda \geq \epsilon/2 > 0$.

2) Node i relays type c packets: In this case, $\mathbb{E}\{A_{ic}(\tau)\} = 0$. Since the 2 Hop Relay Algorithm schedules transmissions out of relay nodes with higher probability than into relay nodes, and using the definition of K, we have:

$$\begin{aligned}\frac{1}{K}\sum_{\tau=t}^{t+K-1} \mathbb{E}\Big\{\sum_b \mu_{ib}^{(c)}(\tau) - \sum_a \mu_{ai}^{(c)}(\tau)\Big\} &\geq r_3 - r_2 - \delta \\ &\geq \frac{2\delta}{1+\delta} r_3 > \epsilon \qquad \text{for} \qquad \delta > \frac{\epsilon}{2r_3 - \epsilon}\end{aligned}$$

Thus, the last summation in (5) for this case is positive as well.

Choosing $\delta = \max(\frac{d\epsilon}{R_1 p + R_2 p' - R_1 q - R_2 q' - d}, \frac{\epsilon}{2r_3 - \epsilon})$, we can apply the Lyapunov Drift Lemma to bound the average packet occupancy. Using Little's Theorem, the average delay per packet satisfies the bound in Theorem 2. □

Discussion, Other Extensions and Implications: The proof of the capacity for the cell-partitioned network is general enough to cover any scheduling restriction structure. From (3), it amounts to:

$$\lambda \leq \frac{1}{2N}\mathbb{E}\Big\{\max_{s\in S}\sum_{c=1}^{C}\Big(Y_c^s(t) + X_{10,c}^s(t) + X_{01,c}^s(t)\Big)\Big\}$$

Therefore the optimal resource allocation is to schedule to maximize $\sum_{c=1}^{C}\Big(Y_c^s(t) + X_{10,c}^s(t) + X_{01,c}^s(t)\Big)$ subject to the scheduling restrictions. For the specific cell-partitioned model considered here, this maximization is achieved by following the preference order of the decision variables as described earlier. This enables us to exactly compute the capacity of the network. It is possible to do the same for extensions to this model involving other constraints. For example, if we include the constraint that a user cannot simultaneously transmit and receive, then the above maximization becomes a maximum-match type problem which is difficult to distribute cell by cell. A possible distributed solution is to define an Aloha like MAC scheme in which each node decides to be either a transmitter or receiver with probability 1/2, and then maximize the above metric.

Our model for the channel state process is idealized in that transmissions are assumed to be error-free. But it should be noted that our technique can be used to get expressions for the capacity region when the channel states on *each* link are described by some fading process which in turn affects the achievable transmission rates. We would like to point out that the capacity will not be a scaled version of the value calculated for the basic model; rather, multi-user diversity can be exploited to maintain competitive capacity region. It is also possible to use our technique for scenarios where each user follows its own distinct mobility process, although the asymmetry may make the expressions for the capacity region more complex.

The 2 Hop Relay Algorithm presented earlier makes scheduling decisions purely based on the current user locations. It certainly does not attempt to minimize the delay in the network. The main objective behind this algorithm is to establish the sufficient condition for the capacity region theorem. The delay performance can be improved using alternative scheduling strategies. In fact, schemes that make use of queue backlog information in making scheduling decisions are shown to outperform a stationary, randomized scheduling strategy in [1]. Similarly, schemes that exploit the mobility pattern of the users (e.g. [9]) can get better delay performance.

IV. Minimum Energy Function

We now investigate the minimum energy function of the cell-partitioned network under consideration. The minimum energy function $\Phi(\lambda)$ is defined as the minimum time average power required to stabilize an input rate λ per user, considering *all* possible scheduling and routing algorithms that conform to the given network structure. In [10], [12], the minimum energy function is formally defined as the solution to an optimization problem. We use the same definition.

We exactly compute this function for our network model. Specifically, we assume that all users receive packets at the same rate (i.e. $\lambda_i = \lambda$ for all i). Also, we consider the case when $R_1 \geq 2R_2$. ($\Phi(\lambda)$ for the case when $R_1 < 2R_2$ has a different expression, but the proof is similar).

Theorem 3: The minimum energy function $\Phi(\lambda)$ per user of the cell-partitioned network is a piecewise linear curve given by the following:

$$\Phi(\lambda) = \begin{cases} \frac{\lambda}{R_1} & \text{if } A \\ \frac{2q}{2d} + \frac{2}{R_1}\left(\lambda - \frac{2R_1 q}{2d}\right) & \text{if } B \\ \frac{(p+q)}{2d} + \frac{1}{R_2}\left(\lambda - \frac{R_1(p+q)}{2d}\right) & \text{if } C \\ \frac{(p+q+2q')}{2d} + \frac{2}{R_2}\left(\lambda - \frac{R_1(p+q)+2R_2q'}{2d}\right) & \text{if } D \end{cases}$$

where
$A \equiv 0 \leq \lambda < \frac{2R_1 q}{2d}$
$B \equiv \frac{2R_1 q}{2d} \leq \lambda < \frac{R_1(p+q)}{2d}$
$C \equiv \frac{R_1(p+q)}{2d} \leq \lambda < \frac{R_1(p+q)+2R_2q'}{2d}$
$D \equiv \frac{R_1(p+q)+2R_2q'}{2d} \leq \lambda < \mu$

Thus, the network can stably support users simultaneously communicating at any rate $\lambda < \mu$ with an energy cost that can be pushed arbitrarily close to $\Phi(\lambda)$ (at the cost of increased delay). We prove the theorem in 2 parts. First, we establish the necessary condition by deriving a lower bound on the energy cost of *any* stabilizing algorithm. Then, we establish sufficiency by presenting a specific scheduling strategy and showing that the average delay is bounded under that policy.

Proof of Necessity: Consider a scheduling strategy that stabilizes the system. Let $X_{ab}(T)$ denote the number of packets delivered by the strategy from sources to destinations in $(0, T)$ that involve exactly a same cell and b adjacent cell transmissions. For simplicity, assume that the strategy is ergodic[3] and yields well defined time average energy expenditure e and well defined time average values for x_{ab} where:

$$x_{ab} \triangleq \lim_{T\to\infty} \frac{1}{T} X_{ab}(T)$$

The average energy cost e of this policy satisfies:

$$e \geq \sum_{a,b} (a/R_1 + b/R_2) x_{ab} \tag{9}$$

This follows by noting that enough packets may not be available during a transmission.

Note that $x_{00} = 0$, and so the only possible non-zero x_{ab} variables are for (a, b) pairs that are integers, non-negative, and such that $(a, b) \neq (0, 0)$. Let $\boldsymbol{x} = (x_{ab})$ represent the collection of x_{ab} variables, and note that these variables must satisfy the constraint $\boldsymbol{x} \in \Omega_0 \cap \Omega_1 \cap \Omega_2 \cap \Omega_3$, where the four constraint sets are defined below:

[3]The general case can be proven similarly.

$$\begin{aligned} \Omega_0 &\triangleq \left\{ \boldsymbol{x} \,\middle|\, x_{00} = 0 \;\&\; \sum_{a,b} x_{ab} = \lambda \right\} \\ \Omega_1 &\triangleq \left\{ \boldsymbol{x} \,\middle|\, \frac{x_{10}}{R_1} \leq c_1 \right\} \\ \Omega_2 &\triangleq \left\{ \boldsymbol{x} \,\middle|\, \frac{1}{R_1} \sum_{a,b} a x_{ab} \leq c_1 + c_2 \right\} \\ \Omega_3 &\triangleq \left\{ \boldsymbol{x} \,\middle|\, \frac{1}{R_1} \sum_{a,b} a x_{ab} + \frac{x_{01}}{R_2} \leq c_1 + c_2 + c_3 \right\} \end{aligned}$$

where c_1 is the maximum rate of source-destination transmission opportunities in the same cell, $c_1 + c_2$ is the maximum rate of all possible same cell transmission opportunities and $c_1 + c_2 + c_3$ is the maximum rate of all same cell or source-destination adjacent cell transmission opportunities (without double counting).

Using the results from Theorem 1, we know that $c_1 = q/d, c_1 + c_2 = p/d, c_1 + c_2 + c_3 = (p + q')/d$. For example, $(p + q')/d$ can be written as $\frac{1}{T}\sum_{t=0}^{T} X_1(t) + X_2(t) + X_3(t)$ where $X_1(t)$ is the maximum number of direct same cell opportunities, $X_2(t)$ is the maximum number of indirect same cell opportunities *given* all direct opportunities are used and $X_3(t)$ is the maximum number of direct adjacent cell opportunities *given* all same cell opportunities are used. Since only one of these three opportunities can used is a given cell in a time slot, the maximum total sum is fixed and hence $c_1 + c_2 + c_3 = (p + q')/d$.

Define $f(\boldsymbol{x}) \triangleq \sum_{a,b}(a/R_1 + b/R_2)x_{ab}$, which is simply the right hand side of (9). Because $e \geq f(\boldsymbol{x})$, and because $\boldsymbol{x} \in \Omega_0 \cap \Omega_1 \cap \Omega_2 \cap \Omega_3$, we have that:

$$e \geq \inf_{\boldsymbol{x} \in \Omega_0 \cap \Omega_1 \cap \Omega_2 \cap \Omega_3} f(\boldsymbol{x}) \tag{10}$$

Furthermore, for any function $g(\boldsymbol{x})$ such that $g(\boldsymbol{x}) \leq f(\boldsymbol{x})$ for all $\boldsymbol{x}$, and for any set $\tilde{\Omega}$ that contains the set $\Omega_0 \cap \Omega_1 \cap \Omega_2 \cap \Omega_3$, we have:

$$e \geq \inf_{\boldsymbol{x} \in \tilde{\Omega}} g(\boldsymbol{x}) \tag{11}$$

This follows because the function to be minimized is smaller, and the infimum is taken over a less restrictive set.

We define four new constraint sets $\tilde{\Omega}_0, \tilde{\Omega}_1, \tilde{\Omega}_2, \tilde{\Omega}_3$ as follows:

$$\begin{aligned} \tilde{\Omega}_0 &\triangleq \Omega_0 \\ \tilde{\Omega}_1 &\triangleq \Omega_1 \\ \tilde{\Omega}_2 &\triangleq \left\{ \boldsymbol{x} \,\middle|\, \frac{x_{10}}{R_1} + \frac{2}{R_1} \sum_{a\geq 2} x_{a0} \leq c_1 + c_2 \right\} \\ \tilde{\Omega}_3 &\triangleq \left\{ \boldsymbol{x} \,\middle|\, \frac{x_{10}}{R_1} + \frac{2}{R_1} \sum_{a\geq 2} x_{a0} + \frac{x_{01}}{R_2} \leq c_1 + c_2 + c_3 \right\} \end{aligned}$$

It can be seen that each of $\Omega_0, \Omega_1, \Omega_2, \Omega_3$ is a subset of $\tilde{\Omega}_0, \tilde{\Omega}_1, \tilde{\Omega}_2, \tilde{\Omega}_3$ respectively. Therefore, $\Omega_0 \cap \Omega_1 \cap \Omega_2 \cap \Omega_3$ is a subset of $\tilde{\Omega}_0 \cap \tilde{\Omega}_1 \cap \tilde{\Omega}_2 \cap \tilde{\Omega}_3$.

We now find four different bounds for e, each having the form $e \geq \alpha\lambda + \beta$. These bounds define the four piecewise linear regions of $\Phi(\lambda)$. Throughout, we assume that $2/R_1 \leq 1/R_2$, so that:

$$\frac{1}{R_1} < \frac{2}{R_1} \leq \frac{1}{R_2} < \frac{2}{R_2} \tag{12}$$

(the opposite case $2/R_1 > 1/R_2$ can be treated similarly, and leads to a different set of bounds).

1) Note that $f(\boldsymbol{x}) \geq \frac{1}{R_1}\sum_{a,b} x_{ab}$. Therefore:

$$e \geq \inf_{\boldsymbol{x}\in\tilde{\Omega}_0} \frac{1}{R_1}\sum_{a,b} x_{ab}$$

Because the constraint $\tilde{\Omega}_0$ is given by $\sum_{a,b} x_{ab} = \lambda$, the above infimum is equal to λ/R_1. Therefore we have our first linear constraint for any algorithm that yields a time average energy of e:

$$e \geq \frac{\lambda}{R_1} \tag{13}$$

2) Next note that:

$$f(\boldsymbol{x}) \geq \frac{x_{10}}{R_1} + \frac{2}{R_1}\sum_{(a,b)\neq(1,0)} x_{ab}$$

This is because $(a/R_1 + b/R_2) \geq 2/R_1$ for any non-negative integer pair (a,b) such that $(a,b) \neq \{(0,0),(1,0)\}$ (using (12). Therefore, we have:

$$e \geq \inf_{\boldsymbol{x}\in\tilde{\Omega}_0\cap\tilde{\Omega}_1} \left[\frac{x_{10}}{R_1} + \frac{2}{R_1}x_{ab}\sum_{(a,b)\neq(1,0)}\right]$$

The right hand side is equal to the solution of the following problem:

$$\begin{array}{ll} \text{Minimize:} & \frac{x_{10}}{R_1} + \frac{2}{R_1}\sum_{(a,b)\neq(1,0)} x_{ab} \\ \text{Subject to:} & \sum_{a,b} x_{ab} = \lambda \\ & \frac{x_{10}}{R_1} \leq c_1 \end{array}$$

The above optimization is equivalent to minimizing $x_{10}/R_1 + 2/R_1(\lambda - x_{10})$ subject to $x_{10}/R_1 \leq c_1$. The solution is clearly to choose $x_{10} = R_1c_1$, and hence we have:

$$e \geq \frac{2\lambda}{R_1} - c_1 \tag{14}$$

3) Next note that:

$$f(\boldsymbol{x}) \geq \frac{x_{10}}{R_1} + \frac{2}{R_1}\sum_{a\geq 2} x_{a0} + \frac{1}{R_2}\sum_{b\neq 0} x_{ab}$$

which follows from case 2 as well as because $1/R_2 \leq b/R_2$ for all positive b.
Thus, we have:

$$e \geq \inf_{\boldsymbol{x}\in\tilde{\Omega}_0\cap\tilde{\Omega}_1\cap\tilde{\Omega}_2} \left[\frac{x_{10}}{R_1} + \frac{2}{R_1}\sum_{a\geq 2} x_{a0} + \frac{1}{R_2}\sum_{b\neq 0} x_{ab}\right]$$

This is equivalent to the following minimization:

$$\begin{array}{ll} \text{Minimize:} & \frac{x_{10}}{R_1} + \frac{2}{R_1}\sum_{a\geq 2} x_{a0} + \\ & \frac{1}{R_2}(\lambda - x_{10} - \sum_{a\geq 2} x_{a0}) \\ \text{Subject to:} & \frac{x_{10}}{R_1} \leq c_1 \\ & \frac{x_{10}}{R_1} + \frac{2}{R_1}\sum_{a\geq 2} x_{a0} \leq c_1 + c_2 \end{array}$$

where we have aggregated the constraint $\sum_{a,b} x_{ab} = \lambda$. Letting $y = \sum_{a\geq 2} x_{a0}$ and simplifying the optimization metric, we find the above optimization is equivalent to:

$$\begin{array}{ll} \text{Minimize:} & -x_{10}(\frac{1}{R_2} - \frac{1}{R_1}) - y(\frac{1}{R_2} - \frac{2}{R_1}) + \frac{\lambda}{R_2} \\ \text{Subject to:} & \frac{x_{10}}{R_1} \leq c_1 \\ & \frac{x_{10}}{R_1} + \frac{2y}{R_1} \leq c_1 + c_2 \end{array}$$

It is not difficult to show that the above optimization is solved when $x_{10} + 2y = R_1(c_1 + c_2)$. We thus have:

$$\begin{array}{ll} \text{Minimize:} & \frac{\lambda}{R_2} - x_{10}(\frac{1}{R_2} - \frac{1}{R_1}) \\ & -(\frac{R_1(c_1+c_2)-x_{10}}{2})(\frac{1}{R_2} - \frac{2}{R_1}) \\ \text{Subject to:} & \frac{x_{10}}{R_1} \leq c_1 \end{array}$$

The coefficient multiplying the x_{10} variable is $-1/2R_1$, and hence the optimal solution is given by $x_{10} = R_1c_1$, yielding:

$$e \geq \frac{\lambda}{R_2} + (c_1 + c_2) - \frac{R_1}{R_2}(c_1 + \frac{c_2}{2}) \tag{15}$$

4) Finally note that

$$f(\boldsymbol{x}) \geq \frac{x_{10}}{R_1} + \frac{2}{R_1}\sum_{a\geq 2} x_{a0} + \frac{x_{01}}{R_2} + \frac{2}{R_2}\sum_{\geq 2} x_{ab}$$

which follows from case 3 as well as because $2/R_2 \leq b/R_2$ for all $b \geq 2$.
Thus, we have:

$$e \geq \inf_{\boldsymbol{x}\in\tilde{\Omega}} \left[\frac{x_{10}}{R_1} + \frac{2}{R_1}\sum_{a\geq 2} x_{a0} + \frac{x_{01}}{R_2} + \frac{2}{R_2}\sum_{b\geq 2} x_{ab}\right]$$

where
$\tilde{\Omega} = \tilde{\Omega}_0 \cap \tilde{\Omega}_1 \cap \tilde{\Omega}_2 \cap \tilde{\Omega}_3$

This is equivalent to the following minimization (using $\sum_{a,b} x_{ab} = \lambda$):

$$\begin{array}{ll} \text{Minimize:} & \frac{x_{10}}{R_1} + \frac{2}{R_1}\sum_{a\geq 2} x_{a0} + \frac{x_{01}}{R_2} + \\ & \frac{2}{R_2}(\lambda - x_{10} - \sum_{a\geq 2} x_{a0} - x_{01}) \\ \text{Subject to:} & \frac{x_{10}}{R_1} \leq c_1 \\ & \frac{x_{10}}{R_1} + \frac{2}{R_1}\sum_{a\geq 2} x_{a0} \leq c_1 + c_2 \\ & \frac{x_{10}}{R_1} + \frac{2}{R_1}\sum_{a\geq 2} x_{a0} + \frac{x_{01}}{R_2} \leq c_1 + c_2 + c_3 \end{array}$$

Again letting $y = \sum_{a\geq 2} x_{a0}$ and simplifying the optimization metric, we find the above optimization is equivalent to:

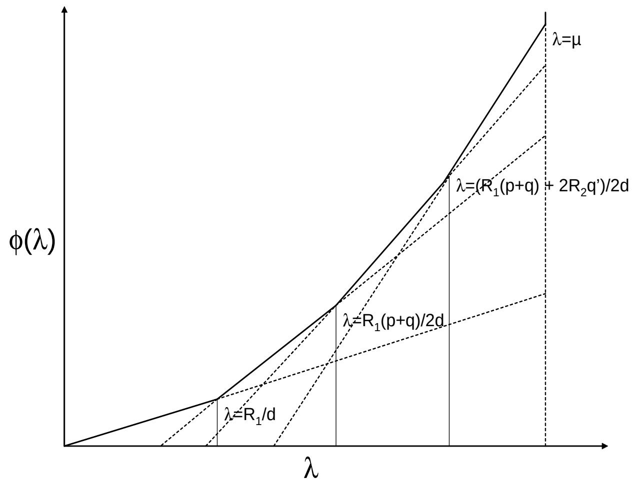

Fig. 2. An illustration of the minimum energy function $\Phi(\lambda)$ showing the piecewise linear structure

$$\begin{aligned} \text{Minimize:} \quad & x_{10}(\tfrac{1}{R_1} - \tfrac{2}{R_2}) + y(\tfrac{2}{R_1} - \tfrac{2}{R_2}) - \tfrac{x_{01}}{R_2} + 2\tfrac{\lambda}{R_2} \\ \text{Subject to:} \quad & \tfrac{x_{10}}{R_1} \leq c_1 \\ & \tfrac{x_{10}}{R_1} + \tfrac{2y}{R_1} \leq c_1 + c_2 \\ & \tfrac{x_{10}}{R_1} + \tfrac{2y}{R_1} + \tfrac{x_{01}}{R_2} \leq c_1 + c_2 + c_3 \end{aligned}$$

The above optimization is solved when $x_{10} = R_1 c_1$, $x_{10} + 2y = R_1(c_1 + c_2)$ and $x_{01} = R_2 c_3$. We thus have:

$$e \geq \frac{2\lambda}{R_2} + (c_1 + c_2) - \frac{R_1}{R_2}(c_1 + 2c_2) - c_3 \qquad (16)$$

Thus, $\Phi(\lambda)$ function is obtained by combining these four bounds. Fig. 2 shows this. We note that $\Phi(\lambda)$ is a piecewise linear function.

Proof of Sufficiency: Now we present an algorithm that makes stationary randomized scheduling decisions and show that for any feasible input rate $\lambda < \mu$, its average energy cost can be pushed arbitrarily close to the minimum value $\Phi(\lambda)$ with bounded delay. However, the delay bound grows asymptotically as the average energy cost is minimized. The exact nature of this energy-delay tradeoff is discussed in the next section.

Similar to the 2 Hop Relay Algorithm, this strategy also restricts packets to at most 2 hops. However, the difference lies in that it *greedily* chooses transmission opportunities involving smaller energy cost over other higher cost opportunities. An opportunity with higher cost is used *only* when the given input rate cannot be supported using all of the low cost opportunities.

Thus, depending on the input rate λ, the algorithm uses only a subset of the transmission opportunities as follows:

1) $0 \leq \lambda < \frac{2R_1 q}{2d}$: In this case, all packets are sent using only source-destination transmission opportunities in the same cell.
2) $\frac{2R_1 q}{2d} \leq \lambda < \frac{R_1(p+q)}{2d}$: Here, all packets are sent either using source-destination transmission opportunities in the same cell or source-relay and relay-destination transmission opportunities in the same cell.
3) $\frac{R_1(p+q)}{2d} \leq \lambda < \frac{R_1(p+q)+2R_2q'}{2d}$: In this case, all packets are sent using same cell or adjacent cell source-destination transmission opportunities.
4) $\frac{R_1(p+q)+2R_2q'}{2d} \leq \lambda < \mu$: Here, all transmission opportunities are used.

To make the presentation simpler, we only discuss the second case where $\frac{R_1 q}{d} \leq \lambda < \frac{R_1(p+q)}{2d}$. The basic idea and performance analysis for other cases are similar. Define a control parameter $V > 1$.

Minimum Energy Algorithm: Every timeslot, for all cells, do the following:

1) If there exists a source-destination pair in the cell, randomly choose such a pair (uniformly over all such pairs in the cell). If the source has new packets for the destination, transmit at rate R_1 with probability $P_{direct} = 1 - \frac{1}{V}$. Else remain idle.
2) If there is no source-destination pair in the cell but there are at least 2 users in the cell, randomly designate one user as the sender and another as the receiver. Then toss an *unfair* coin such that the probability of "Head" is $\frac{1-\delta}{2} P_{relay}$, where $\delta = \delta(\epsilon)$ and $0 < \delta < 1$ and

$$P_{relay} = \frac{\frac{\lambda}{R_1} - \frac{q}{d}(1 - \frac{1}{V})}{\frac{p-q}{2d}}$$

If the outcome is "Head", perform the first action below. Else, perform the second.

a) *Send new Relay packets in same cell:* If the transmitter has at least R_1 packets of type "relay" for its destination, transmit at rate R_1. Else remain idle.
b) *Send Relay packets to their Destination in same cell:* If the transmitter has at least R_1 packets of type "relay" for the receiver, transmit at rate R_1. Else remain idle.

Note that the above algorithm does not use any adjacent cell transmission opportunities. All packets are sent over at most 2 hops using only same cell transmissions. We now analyze the performance of this algorithm.

Theorem 4: For a cell partitioned network (with N nodes and C cells) as described above, with a capacity μ and minimum energy function $\Phi(\lambda)$, the average energy cost e of the Minimum Energy Algorithm with input rates λ for each user such that $\lambda + \epsilon \leq \mu$ for some $\epsilon > 0$ and a control parameter $V = V(\epsilon) > 1$, satisfies:

$$e \geq \Phi(\lambda) + O(\frac{1}{V})$$

with an average packet delay $\overline{D}$ given by:

$$\overline{D} \leq \frac{O(KNV)}{\epsilon}$$

where K represents a parameter indicating the time required for the mobility process to reach a "near steady state", described in more detail in [1].

Proof: It can be seen that for the case under consideration, each user either transmits directly to its destination or transmits new packets to a relay or transmits relayed packets to their destination. Each such transmission involves unit energy cost and therefore the average energy cost e can be expressed in terms of the rates of these transmission opportunities. Let these be r_1, r_2 and r_3 respectively. We have (using results from Theorem 2):

$$
\begin{aligned}
e &= r_1 + r_2 + r_3 \\
&= \frac{q}{d}P_{direct} + \frac{p-q}{d}\frac{1-\delta}{2}P_{relay} + \frac{p-q}{d}\frac{1+\delta}{2}P_{relay} \\
&= \frac{q}{d}P_{direct} + \frac{p-q}{d}P_{relay} \\
&= \frac{q}{d}(1-\frac{1}{V}) + \frac{p-q}{d}\frac{\frac{\lambda}{R_1} - \frac{q}{d}(1-\frac{1}{V})}{\frac{p-q}{2d}} \\
&= \frac{2\lambda}{R_1} - \frac{q}{d}(1-\frac{1}{V}) \\
&= \Phi(\lambda) + \frac{q}{dV}
\end{aligned}
$$

The delay of this scheme can be analyzed using the same procedure used in the proof of Theorem 2. We first apply the Lyapunov Drift Lemma to bound the average packet occupancy. Using Little's Theorem, the average delay per packet satisfies the bound in Theorem 4. □

A natural question to ask is if it is possible to get better delay performance by using alternative scheduling strategies. We discuss this in the next section.

V. Optimal Energy Delay Tradeoff

The optimal energy-delay tradeoff for a single queue over a single fading channel was first characterized in [11]. It was shown that, under strict convexity assumptions on the rate-power curve of the system, any set of algorithms (parameterized by $V > 1$) that yield average power required for stability within $O(1/V)$ of the minimum power required for stability must have average queueing delay greater than or equal to $\Omega(\sqrt{V})$. It has recently been shown in [12] that for multi-user wireless downlink systems, if the minimum energy function $\Phi(\vec{\lambda})$ is piecewise linear about the rate vector $\vec{\lambda}$, then it is possible to beat this $\Omega(\sqrt{V})$ bound. In fact, a delay of $O(\log(V))$ is achievable as the average power is pushed within $O(1/V)$ of the minimum power, although this was only shown for single hop networks.

Since the minimum energy function for the cell-partitioned network that we consider here has a piecewise liner structure, it would be interesting to develop such an algorithm for this network that can achieve the optimal energy-delay tradeoff.

VI. Conclusions

In this work, we investigated two quantities of fundamental interest in a mobile ad-hoc network: the network capacity and the minimum energy function. Using a cell-partitioned model of the network, we obtained *exact* expressions for both these quantities in terms of the network parameters (number of nodes N and number of cells C) and the steady state location distribution of the mobility process. Our results hold for general mobility processes (possible non-uniform and non-i.i.d) and our analytical technique can be extended to other models with additional scheduling constraints.

We also proposed scheduling strategies that can achieve these bounds arbitrarily closely at the cost of an increased delay. The key observation is that such capacity and minimum energy achieving schemes are necessarily 2-hop relay algorithms. Although these schemes make scheduling decisions purely based on the current user location, their average delay is bounded. The delay performance can be improved using other information (like queue backlogs, mobility patterns etc.).

We have focused on network control algorithms that operate according to the given network structure. We assumed that the packets themselves are kept intact and are not combined or network coded. An interesting future direction of this research would be to discover if network coding can be used to either increase capacity or reduce average energy expenditure, perhaps by augmenting the network model to include coding options and to also exploit the broadcast channel structure, as in [13].

References

[1] M. J. Neely. *Dynamic Power Allocation and Routing for Satellite and Wireless Networks with Time Varying Channels*. PhD Thesis, MIT, LIDS, 2003.

[2] M. J. Neely and E. Modiano. Capacity and Delay Tradeoffs for Ad-Hoc Mobile Networks. *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 1917-1937, June 2005.

[3] P. Gupta and P.R. Kumar. The Capacity of Wireless Networks. *IEEE Transactions on Information Theory*, Vol. 46:388-404, March 2000.

[4] M. Grossglauser and D. Tse. Mobility increases the Capacity of Ad-Hoc Wireless Networks. *Proceedings of IEEE INFOCOM*, April 2001.

[5] G. Mergen and L. Tong. Stability and Capacity of Regular Wireless Networks. *IEEE Transactions on Information Theory*, Vol. 51, Issue 6, June 2005 Page(s):1938 - 1953

[6] S. Toumpis and A. J. Goldsmith. Large wireless networks under fading, mobility, and delay constraints. *Proceedings of IEEE INFOCOM*, March 2004.

[7] A. El Gammal, J. Mammen, B. Prabhakar, and D. Shah. Throughput-delay trade-off in wireless networks. *Proceedings of IEEE INFOCOM*, March 2004.

[8] X. Lin and N. B. Shroff. The fundamental capacity-delay tradeoff in large mobile ad hoc networks. *Purdue University Tech. Report*, 2004.

[9] M. Grossglauser and M. Vetterli. Locating Nodes with EASE. *Proceedings of IEEE INFOCOM*, April 2003.

[10] M. J. Neely. Energy Optimal Control for Time Varying Wireless Networks. *Proceedings of IEEE INFOCOM*, March 2005.

[11] R. Berry and R. Gallager. Communication over Fading Channels with Delay Constraints. *IEEE Transactions on Information Theory*, vol. 48, no. 5, pp. 1135-1149, May 2002.

[12] M. J. Neely. Optimal Energy and Delay Tradeoffs for Multi-User Wireless Downlinks. *USC Technical Report CSI-05-06-01*, June 2005.

[13] Y. Wu, P. A. Chou, and S.-Y. Kung. Minimum-energy multicast in mobile ad hoc networks using network coding. to appear in *IEEE Transactions on Communications*