Spectrum and Energy Efficiency of Mobility-Assisted Content Distribution in Cellular Networks

Vladimir Vukadinovic^{*} and Gunnar Karlsson[†]

^{*}Disney Research Zurich
vvuk@disneyresearch.com

[†]School of Electrical Engineering and ACCESS Linnaeus Centre
KTH - Royal Institute of Technology
gk@ee.kth.se

Abstract— We consider a wireless content distribution service for mobile users that can be provided by mobile operators in their coverage areas. Our focus is on dense urban environments. A basic assumption is that the latency and throughput requirements of the service are flexible (delay tolerance) and the contents thus do not need to be delivered to all users via the operator's network: They can be relayed from one user to another in a store-carry-and-forward fashion. Hence, mobility of users can be used as a supplementary transport mechanism to off-load the operator's network and relieve strain on scarce spectrum resources. In this paper, we evaluate the achievable throughputs and spectrum savings of such a mobility-assisted wireless architecture for various mobility patterns and user densities in an area. We also estimate the energy consumption of the mobility-assisted content distribution for the case when Bluetooth radios are used for peer-to-peer communication between users. Our goal is to determine whether mobility assistance brings any advantages for delivering delay-tolerant contents in a wireless network.

Index Terms— Mobile communication systems, multicast, content distribution, portable devices, store and forward networks.

1 INTRODUCTION

Continuous wireless connectivity in mobile communications requires coverage that is increasingly more difficult to provide due to expectations of higher bit rates while there are limits in spectrum and power as well as in economical resources of network operators. We start herein from the premise that continuous connectivity is not needed for all services and hence intermittent communication may be useful for applications characterized by a low degree of interactivity, e.g., broadcasting, paging, messaging, and data collection. We consider an architecture that consists of a cellular network supplemented by a communication mode that relies on relaying of data among mobile nodes. The nodes communicate with one another when they are within radio range and carry the data onwards through their own movements. The mobility patterns of nodes affect the speed, throughput and reliability of the data forwarding.

The content distribution is provided by multicast from the cellular network and the data is further disseminated by the mobile nodes themselves; the service is best effort and does not give any latency guarantees. Some users receive the contents from base stations using licensed spectrum. They will carry the contents onwards through their own movements and deliver them to others using short-range radios that operate in license-free spectrum. Contents suitable for the service are those that to-

day are distributed by podcasting, as well as updates of software, timetables and maps, or similar bulk data transfers to groups of users.

The mobility-assisted content distribution could reduce the spectrum usage in cellular networks by transmitting the contents only to the users that enjoy favorable radio conditions. The throughput of the system is however limited by the ability of the users to forward the contents to others, from node to node. This ability strongly depends on human mobility characteristics—we are primarily concerned with pedestrian mobility herein. We study the achievable throughput and spectrum savings of mobility-assisted content distribution for various mobility scenarios and users densities. We also study the energy consumed in the mobile devices to receive and forward the contents. Our study shows the following: i) Leveraging delay tolerance, to allow service outages, is an economical alternative to deploying denser cellular infrastructures. ii) Large savings in licensed spectrum usage are possible with an architecture that uses both infrastructure and ad hoc communication. iii) The throughput of the mobility-assisted content distribution increases with the density of users, which is opposite of the scaling in infrastructure networks. iv) Depending on the mobility, density, and the extent of cooperation among users, the mobility-assisted content distribution may or may not be more energy efficient than the infrastructure-based service.

Our study is based on simulations. The complexity of the considered system prevents analytical modeling. For example, pedestrian mobility, which strongly affects the performance, is not mathematically tractable. Our simulation model is however rather general. It abstracts away the details of networking protocols and technologies used in the mobile devices (see [1] for a system description). Furthermore, the study is relevant for any delay-tolerant service based on a publish/subscribe paradigm (see [2] for an introduction to pub/sub).

Our work belongs to the general area of delay-tolerant networking, which aims at providing communications in performance-challenged environments characterized by high degrees of intermittency [3]. Overviews of challenges in the area can be found in [4][5]. The performance of unicast forwarding algorithms in such networks has been studied extensively. Based on the mobility traces collected in the Huggle project [6], researchers have studied the characteristics of human mobility and how it impacts message forwarding [7]. The mobility-assisted content distribution considered here is a multicast/broadcast service, which relies on epidemic forwarding. Applications of gossip-based epidemic forwarding to mobile communication have been proposed in, for instance, [14] – [18]. A peer-to-peer file sharing application for Bluetooth enabled devices has been described in [19]. Multicast for delay-tolerant networks has been proposed in [20]. The impact of user cooperation on the efficiency of content distribution in intermittently connected networks has been studied in [21]. The type of service that we considered here is best described in [22],[23]. We study how such architecture can be used to sup-

plement cellular infrastructures. A concise version of this paper has been published in [24]¹.

The impact of introducing islands of connectivity provided by fixed infrastructures (relays, access points, stand-alone caching nodes, et cetera) on the performance of delay-tolerant ad-hoc networks has been addressed in [8] – [10]. For example, [10] considers a “data push” service in a hybrid WLAN network and compares throughput contributions of access points and opportunistic contacts. The spectrum usage has not been addressed. In [8] – [10], deployments of sparse infrastructures are studied as a means to enhance the performance of ad-hoc networks. Our scenarios are similar to the one presented in [11]: The authors propose cooperative web browsing where collocated devices download complementary parts of a web page from a base station and then exchange them via Bluetooth. Impact of mobility however has not been considered: cooperating nodes are assumed to be static. Works in [8] – [11] provide user-centric views by focusing on performances relevant to end users. The benefits/incentives for network operators to allow and support direct mobile-to-mobile communication are not sufficiently well understood. Our work provides an operator-centric view by focusing on the possible improvements in spectral-efficiency of their cellular systems. In that sense, our work is related to [12], [13], where authors investigate how WLAN infrastructure (e.g. Wi-Fi hot spots) can be used to off-load cellular networks. Our focus is on the use of ad-hoc contacts among mobile devices: no fixed WLAN infrastructure is assumed or needed to support mobility-assisted podcasting.

The structure of the paper is as follows. The spectral efficiency of infrastructure-based multicasting for content distribution in a cellular network is studied in Section 2. In Section 3, we describe the supplementary mobility-assisted content distribution scheme in more details and introduce our evaluation methodology. The performance results for the achievable throughput and spectrum usage are presented in Section 4. The energy consumption of mobility-assisted scheme is addressed in Section 5. We conclude our findings in Section 6.

2 INFRASTRUCTURE-BASED PODCASTING

In this section, we study the licensed-spectrum usage of the infrastructure-based content distribution. In later sections, these results are compared with those obtained for the mobility-assisted distribution. The setup is similar to the one described in [25]: We consider a 10×10 km² area covered by a hexagonal grid of base stations. The area is wrapped around to form a torus to avoid boundary effects. The cell radius is $R = 1$ km in the baseline scenario (~ 0.38 base stations per km²). There are 32 orthogonal sub-channels in each cell. The density of users in the area may vary from 10 to 50 users per km². The users are subscribed to the same content feed and form multicast groups in corresponding cells. In addition,

¹ We now include a case study on energy efficiency and improve the study of spectrum usage by considering different types of cooperation among users and a broader set of network parameters.

there are 500 non-subscribed users in the area that occupy some of the sub-channels and, therefore, create inter-cell interference for the subscribers. The propagation model depends on the distance of a user from the base station: For a user that is close to its base station, we assume line of sight propagation with the loss exponent 2. For a user that is far away (no line of sight), we assume the Hata propagation loss model [26] with the loss exponent 3.52 and zero-mean Gauss-distributed shadowing with standard deviation 8 dB. For each user/base-station pair, the transition distance between the line of sight and no line of sight is randomly chosen between 30 m and 70 m. One by one, the non-subscribed and the subscribed users are randomly placed into the area and associated with a base station to which they have the strongest path. Each user is allocated a sub-channel on which he experiences the lowest interference from the users added before him. Since the contents are delivered in multicast mode, all subscribed users in a cell are allocated the same sub-channel, which is determined at the time when the first of them is added to the cell. We assume that there is no power control in the system: all base-station transmits at the same power on each occupied sub-channel. All nodes are stationary. The total transmission power per unit area is a parameter that we vary. We calculate the signal to noise-plus-interference ratio (γ) for each user and assume that his achievable rate is 3 dB from the Shannon limit (hence, the spectral efficiency is given by $\log_2(1 + \gamma/2)$). This is a reasonable assumption in modern systems, such as LTE [27]. In our baseline scenario, we assume that all subscribed users should be able to receive the contents. Therefore, the multicast transmission rate in a cell is limited by the achievable rate of the weakest subscriber in that cell.

In Fig. 1 we show the spectral efficiency and spectrum usage for 50 subscribed users per km^2 and various power budgets. The results are averaged over 100 simulation runs; the average coefficient of variation between runs is approximately 3 %. The spectrum usage in the area is calculated as the reciprocal of the spectral efficiency multiplied by the ratio of the cells that have at least one subscribed user (because spectrum is not used in cells without subscribers). In our baseline scenario, this ratio is equal to one with high probability due to the low density of base stations. Therefore, the spectrum usage shown in Fig. 1 is simply a reciprocal of the spectrum efficiency, which will not be the case in scenarios where we vary the base station density. The results show that the spectral efficiency increases with the power as long as the system is noise-limited—it levels off when the system becomes interference-limited. The spectrum usage in the interference-limited regime is around 2 Hz/b/s. In the following, we consider two possible strategies to reduce the spectrum usage.

The first strategy is to deploy a denser infrastructure by adding more base stations. The benefits of the denser deployments are higher spectral efficiency (base stations are located closer to the users) and better localization of the used spectrum (number of “empty” cells increases). The second strategy is to relax the assumption that contents are transmitted at the achievable rate of the weakest subscriber in

Fig. 1. Average spectral efficiency and spectrum usage for 50 subscribers per km^2 for various power budgets.

the cell. By selecting a higher transmission rate, some of the subscribers will be in *outage*. The spectral efficiency increases if we allow more subscribers to be in outage so that only the strongest are served. We assume that outaged subscribers will use their short-range radios to connect to other users to download the missing contents. We evaluate the feasibility of this assumption in Section 4; here we focus on the performance of the cellular network only. The two strategies are expected to provide similar results with respect to the spectral efficiency since dropping the weak subscribers effectively shrinks the coverage areas of the cells.

The spectral efficiency and spectrum usage are shown in Fig. 2 for various base station densities and various percentages of served subscribers p (the outage probability is $1 - p$). In all scenarios, the transmission power budget is set to a value large enough to ensure that the network is in the interference-limited regime. When varying the base station density, we set the heights of base station antennas to $6.5 \times \log(R/10)$ (from 30 m for $R = 1000$ m down to 15 m for $R = 100$ m). When varying p , we assume that the p -percent of the strongest subscribers in the area are selected and the transmission rate in each cell is chosen based on the achievable rate of the weakest among the selected subscribers in that cell. As expected, the spectral efficiency increases with the density of the base stations and with the number

Fig. 2. Spectral efficiency (left) and spectrum usage (right): The base station density and the equivalent cell radius are both indicated on horizontal axes located below and above each plot, respectively.

of subscribers in outage (i.e. with the decrease in the service probability p). It is limited from above by the spectral efficiency of a subscriber located at the base of a cell tower. The spectrum usage depends on the spectral efficiency and the number of occupied cells (i.e. cells with at least one content subscriber), as described earlier. Both when the cell density increases and the service probability p decreases, the spectrum usage is significantly reduced. This is due to the increase in spectral efficiency and due to the appearance of spatial spectrum holes. The results indicate that, for example, by allowing 50% of the subscribers to be in outage the spectrum usage can be reduced as much as by deploying up to 10 times denser infrastructure. This is however possible only if the outaged subscribers can receive the missing contents through p2p contacts with other users. Hence, there is a strong incentive to explore the feasibility of mobility-assisted content distribution, considering that the deployment of a denser infrastructure can be prohibitively expensive.

3 MOBILITY-ASSISTED CONTENT DISTRIBUTION

We present a simplified model of mobility-assisted content distribution to study the achievable throughput, licensed spectrum usage, and energy-efficiency. We consider a group of users subscribed to the same content feed – we refer to them as *subscribers*. New contents are continuously published on the feed and transmitted from a group of base stations at a constant rate T_{BS} . At any time, p -percent of the subscribers with the strongest propagation channels towards the base stations are selected to receive contents (Fig. 3). This selection criterion minimizes the licensed spectrum usage in the network in accordance with Section 2. However, we now have mobility of users. Depending on the rate of channel fluctuations, the set of selected subscribers may change rapidly, or remain constant for prolonged periods of time. The rate of channel fluctuations depends on the mobility of subscribers and the propagation environment (i.e. shadowing and multipath fading). In our simulations, we do not model the details of the propagation environment. Therefore, in Section 3.2, we introduce some simplifying methods to decide on the set of selected subscribers.

The selected subscribers forward contents to other subscribers using short-range radios (SRRs), when contact opportunities arise. The contents can also be forwarded to users that are not subscribed to the feed, but who are willing to ferry the contents on behalf of the subscribers – we refer to them as *supporters*. Hence, we assume that users subscribed to different feeds may cooperate in spreading the contents (but we study only the performance of one of those feeds). The supporters give preemptive priority to the contents to which they are subscribed. Only when they are not busy forwarding their own contents will they ferry contents for the subscribers. We use term *user* to refer to any user: a subscriber or a supporter. The density of the subscribers reflects the popularity of the feed, while the density of the users reflects the popularity of the content distribution service. The ratio of the two densities

gives the relative popularity of the feed among the users.

We assume that the mobile devices continuously search for contact opportunities and instantaneously establish connections via SRRs when they discover each other. Our interest is in the achievable performance of the mobility-assisted distribution (throughput, licensed spectrum usage and energy consumption), unrestricted by the limitations of current SRR technologies. We are concerned with fundamental constraints that cannot be engineered – those constraints stem from the density and mobility of the users. Therefore, we do not make assumptions on any particular SRR technology to be used. We consider two simple models of the physical layer in which users exchange data at the constant rate offered by their SRRs if they are within a *fixed* transmission range of one another:

- The *multi-channel* model assumes that the unlicensed frequency band used by the SRRs is divided into many channels. Transmitters employ fast channel hopping to avoid interference. The number of channels ensures that there is an interference-free channel at any time. Connections between users are full-duplex/TDD and strictly pairwise. This model resembles Bluetooth radio.
- The *single-channel* model assumes that the unlicensed frequency band used by the SRRs constitutes a single channel. When a user starts to transmit his data, no other data can be received or transmitted in the interference regions (disks) around the transmitter and the intended receiver(s), respectively. A user may transmit to several users at a time (i.e. broadcast). Connections are half-duplex. This model resembles Wi-Fi (there are multiple channels in Wi-Fi, but few of them are non-overlapping).

The main difference between the two models is in the way they handle possible interference. In the multi-channel model, we assume that the number of non-interfering channels N_C is large enough to avoid interference for any considered user density. This however means that the nominal data rate offered by a multi-channel SRR is N_C times lower than the rate of a single-channel radio. We do not aim

Fig. 3. Mobility-assisted podcasting: Circular nodes are users subscribed to the same content feed. Some of them are selected to receive contents from base stations (darkened) and forward them to other nodes. Square nodes are users subscribed to other feeds, which may support the forwarding.

to provide a quantitative comparison of performances achieved with the two SRR models. Therefore, all throughput results presented in the following sections are *normalized* by the nominal data rates of SRRs.

The exchange of contents between devices proceeds as follows: When two devices enter each other's range, one of them (decided randomly) will initiate a connection and request the list of available contents on the second device. Each byte of data is indexed so that users can identify and exchange missing data. Users always give preference to the lower indexes when they download the missing data, trying to fill the gaps in the sequence of bytes. We assume that every byte of data transferred between two users is meaningful for the application. There is no notion of content size in our setup: the impact of incomplete transfers is ignored. In other words, each byte can be considered as a content item on its own. With the (half-duplex) single-channel model, the second device waits until the initiating device finishes its turn before it starts downloading missing data. With the (full-duplex/TDD) multi-channel model, both devices may exchange missing data simultaneously, but at the rate equal to one-half of the nominal data rate offered by the SRR. The connection may break at any time due to mobility. While connected to each other, the two devices are oblivious to the requests of other devices around them. With the single-channel model, we however allow neighboring devices to overhear the data exchange between the two devices and obtain missing data. This is what we refer to as "broadcasting".

3.1 Mobility Scenarios

The performance of mobility-assisted content distribution depends on the rate and the duration of contact opportunities. Therefore, capturing users' mobility is an important part of our study. We consider cases where users move according to three different mobility scenarios.

Truncated Levy Walk (TLW): In this scenario, users are moving according to the TLW model proposed in [28]. The TLW model describes a random walk defined by a sequence of steps that users make. Each step represents a trip ("flight") followed by a pause. At the beginning of each step, a user chooses a random direction/angle from a uniform distribution and his flight length and pause time from truncated power-law distributions. The model is derived from mobility-tracking logs obtained from volunteers carrying GPS receivers in various outdoor settings [28]. The authors observed that the recorded mobility patterns have features defining Levy walks: flight length and pause time distributions closely match truncated power-law distributions. We assume that the flight lengths and pause times follow Cauchy distributions with scale parameters equal to one. The flight lengths are truncated to [10, 1000] m. The lengths are not restricted by distances from the edges of the simulated area because the area is wrapped around to a torus. Pause times are truncated to [10, 3600] s. For each flight,

Fig. 4. A part of Chicago's downtown area.

users choose speeds uniformly from $[0.5, 2]$ m/s.

Constrained Random Waypoint (CRWP): In this scenario, users are moving according to the well-known random waypoint model on a topology that represents a part of Chicago's downtown area (Fig. 4). In each step, users choose random waypoints in the area. Users are restricted to travel between waypoints using sidewalks and walkways. Pause times at the waypoints and traveling speeds are chosen from the same distributions as in the previous scenario.

UDEL Mobility Model (UDEL): In this scenario, we distinguish between "non-working" and "working" users, which comprise respectively 30% and 70% of the total population. The non-working users (e.g. tourists, shoppers) move according to the described constrained random waypoint model. The working users move according to an activity model that was developed based on empirical data from the US department of labor statistics [29][30]. The buildings shown in Fig. 4 can be residential, or they can contain offices and stores. The working users initially appear at random residential locations (homes), and move according to the activity model: They commute between their homes and offices and possibly between offices and other locations during lunch breaks. Their walking speeds are drawn uniformly from $[0.5, 2]$ m/s as in the previous scenarios. However, they are stationary most of the time (sitting in offices). More details are provided in [29][30]. We simulated different periods of the day to capture different parts of the working day cycle.

Some examples of users' trajectories for a one-hour period are shown in Fig. 5. The patterns produced by the TLW model may, for example, represent the movements of pedestrians in theme parks, state fairs, and university campuses. Tourists and shoppers in city areas are better represented by the CRWP model. For the same user density, the CRWP model creates more contact opportunities than the TLW model due to the space-constrained movements. The UDEL model is perhaps closest to urban pedestrian mobility encountered in reality. During a random one-hour period, the mobility of most

Fig. 5. Examples of nodes' trajectories for TLW (left), CRWP (middle), and UDEL (right) mobility models.

people is confined to their homes, offices, classrooms and other spaces where they perform their daily activities. There is a high risk that such users will have no opportunity to connect to other users. The risk decreases with the user density (popularity of the service) and the number of non-working users.

3.2 Evaluation Method

We describe the method we use to evaluate the throughput and spectrum usage of the mobility-assisted content distribution. Let T_i be the time-average throughput of subscriber i . Since the subscriber may receive contents both through the cellular infrastructure and through the p2p contacts, his throughput has two components: $T_{CELL,i}$ and $T_{P2P,i}$. Hence, $T_i = T_{CELL,i} + T_{P2P,i}$. The throughput depends on how often the subscriber i is selected to receive contents from the base stations and how often he encounters users with missing data. Throughputs of different subscribers may differ significantly. We define the *throughput of mobility assisted content distribution* (T) as the time-average throughput of the subscriber that is least likely to be selected by the base stations ($T = T_i$ where $i = \operatorname{argmin}_j T_{CELL,j}$). For example, if there is a subscriber i who never is selected due to poor channel conditions ($T_{CELL,i} = 0$), the throughput T is then given by his time average peer-to-peer throughput $T_{P2P,i}$. If there is a group of users that are equally least likely to be selected, T is calculated as an average over the group.

As discussed earlier in this section, subscribers are selected based on the quality of their radio channels, which varies over time. Therefore, the subset of selected subscribers may be different in each time-slot. The rate of channel fluctuations depends on the propagation environment and the mobility of subscribers. In dense urban areas, distance to a base station may have little bearing on the channel quality due to intensive shadowing and multipath fading. To simplify the evaluation and make it as general as possible, we *decouple* channel fluctuations from mobility and consider two limiting cases that we refer to as *quasi-static* and *fluid* regimes. The quasi-static regime assumes that the channel fluctuations are so slow that the same subset of subscribers is selected in each time-slot. The fluid regime assumes that the channel fluctuations are so fast that the subset of strongest subscribers becomes random in each time-slot (subscribers are selected with equal probabilities).

In the quasi-static regime, there is a clear distinction between subscribers who always receive contents from the cellular network and subscribers who must always rely on p2p contacts to obtain the contents. Let S be the set of all subscribers and $S' \subset S$ be the subset of selected subscribers. Then, $T_{p2p,i} = 0$ if $i \in S'$ and $T_{CELL,i} = 0$ if $i \in S \setminus S'$. The throughput of the mobility-assisted service T is given by $T = \overline{T_i} = \overline{T_{p2p,i}}$ for $i \in S \setminus S'$ (the bar denotes an ensemble average over i). In the fluid regime, each subscriber is occasionally scheduled to receive contents from the cellular network. The throughput is given by $T = \overline{T_i}$ for $i \in S$.

Performance-wise the main difference between the two regimes is in the following: In the fluid regime, contents are randomly injected into the p2p network throughout the area. The content diversity among the users is higher compared to the quasi-static regime. This reduces the probability that users have the same data and, therefore, have nothing to exchange when they meet each other. Our conjecture is that the quasi-static and fluid regimes give, respectively, lower and upper bounds on the achievable throughput T .

In the following sections, our objective is to evaluate both the throughput T and the content publishing rate from the base stations T_{BS} that achieves T . The rate T_{BS} determines the licensed spectrum usage in the cellular network. When T_{BS} is very low, the capacity of the p2p network (which depends on the mobility of the users and the data rate of their SRR interfaces) is most likely sufficient to deliver all contents to all subscribers. Hence, $T = T_{BS}$. On the contrary, when T_{BS} is high, a large number of contents are injected into the selected users of the p2p network, but few of the contents are delivered to the subscribers that rely on p2p contacts. Hence, spectrum in the cellular network is wasted.

We are concerned with the maximum achievable throughput T . In order to estimate it by simulations, the transmission rate T_{BS} should be carefully set to provide a good balance between the spectrum usage, p2p throughput utilization, and content deliverability. Hence, T_{BS} is considered, not as an input parameter that is varied independently, but as a parameter that needs to be determined in order to estimate T . We rely on a two-stage heuristic to achieve this:

- In the first stage, called probing stage, the base stations transmit contents at a probing rate T_{PB} , which is high enough to fully utilize the p2p network. The maximum p2p throughputs $T_{p2p,i}$ are estimated and T_{BS} is set accordingly. Steps to determine the T_{PB} and T_{BS} in the quasi-static and fluid regimes are provided in the Appendix.
- In the second stage, the base stations transmit contents at the rate T_{BS} . The throughput and spectrum usage of the mobility-assisted content distribution may then be determined. The spectrum usage is compared to that of the infrastructure-based content distribution for the same throughput T . The ratio of the two spectrum usages defines the spectrum saving of the mobility-assisted service.

4 PERFORMANCE RESULTS

We present the results for the throughput and spectrum usage of the mobility-assisted content distribution for various densities of subscribers (from 10 to 50 subscribers per km^2), percentages of subscribers that receive contents from the cellular network p (from 5% to 50%), and overall densities of users (from 100 to 2000 users per km^2). The transmission range of the SRRs is 10 meters unless stated otherwise.

4.1 Impact of Mobility Patterns

In this scenario, we assume that the supporters do not have contents of their own to receive and, therefore, they are willing to forward the contents on behalf of the subscribers at any time. We study the effect of limited cooperation in Section 4.3.

Results obtained with 100 users per km^2 , out of which 50% are subscribed to the feed (50 subscribers plus 50 supporters per km^2), are shown in Fig. 6. The results are averaged over 1000 simulation runs; simulated time in each run was one hour excluding the warm-up period. The limits of 95 % confidence intervals for the throughput T are within 12 % of the shown values. This applies to UDEL scenario in Fig. 6 (bottom left). In all other scenarios, with larger numbers of users or with larger p , the intervals are tighter.

The rows of the figure correspond (top to bottom) to TLW, CRWP, and UDEL mobility scenarios. The left column shows the lower bounds (dashed curves) and upper bounds (solid curves) on the throughput T , which are obtained in the quasi-static and fluid regimes, respectively. The throughput is expressed as a percentage of the data rate offered by the SRR interfaces—the nominal rate offered by the single-channel radio would typically be significantly higher than the rate of the multi-channel radio. In the fluid regime, the difference in throughputs achieved with single- and multi-channel radios is mainly due to the full-duplex capability assumed for the latter. In the quasi-static regime, however, the bytes are downloaded sequentially—when two users connect to each other, only the one falling behind in the sequence will receive the data. Therefore, the full-duplex capability of multi-channel radios does not improve the throughput.

The results indicate that the throughputs in the TLW and UDEL mobility scenarios are in the order of few percents ($< 10\%$) of the SRR interface rate. With slow radios (e.g. Bluetooth versions up to 2.1), the obtained throughputs significantly restrict the size/type of contents that can be delivered to subscribers. This might not be the case with the Wi-Fi and the latest Bluetooth standard (v3.0 is expected to support up to 24 Mb/s). An alternative to increasing the rate of SRRs is to increase the transmission range, which would however increase the energy consumption. The throughputs achieved in the CRWP scenario are significantly higher due to the higher frequency of contacts among users.

Fig. 6. Results with 100 users per km^2 : throughput T (left), average transmission rate from the base stations T_{BS} required to achieve the throughput T (middle), and spectrum usage relative to the infrastructure-based service (right). Different rows correspond to different mobility models: TLW (top), CRWP (middle), and UDEL (bottom). Dashed and solid curves correspond to the quasi-static and fluid regimes, respectively.

The average transmission rate from base stations T_{BS} that achieves the throughput T is shown in the middle column of Fig. 6. T_{BS} determines the average spectrum usage in the cellular network. It increases with the variance of p2p throughputs $T_{p2p,i}$ among subscribers, as discussed in the Appendix. High variability of $T_{p2p,i}$ among subscribers makes it difficult to balance spectrum efficiency and delivery ratio. The spectrum efficiency is measured by the discrepancy between T_{BS} and T . When p2p throughputs of all subscribers are equal, this discrepancy is minimized and spectrum efficiency is maximized. This however cannot be achieved because the variance of p2p throughputs depends on mobility, which cannot be engineered. The largest variance (not shown) is observed in the UDEL scenario, followed by the TLW and CRWP scenarios, respectively. Therefore, the largest discrepancy between T_{BS} and T is in the UDEL scenario. The variance (hence T_{BS}) may decrease with the density of users in the area, as shown later.

In the right column of Fig. 6, we compare the spectrum usages of infrastructure-based and mobility-

assisted content distribution. First, we calculate the spectrum required to provide the transmission rates T_{BS} for various percentages p . The required spectrum is calculated based on the spectrum usage shown in Fig. 2 (right) for $R = 1000$ m. This is the spectrum usage of the mobility-assisted service. Second, we calculate the spectrum required to provide the throughput T via base stations only. The required spectrum is calculated based on the spectrum usage shown in Fig. 2 (right) for $p = 100\%$ and $R = 1000$ m. This is the spectrum usage of the infrastructure-based service. The ratio of these two spectrum usages is shown in the right column of Fig. 6.

For the same throughput T , the mobility-assisted service uses only a fraction of the spectrum required for the infrastructure-based service, as shown in Fig. 6. The largest spectrum savings (hence the lowest spectrum usage ratios) are achieved when only a small subset of subscribers is served via base stations. However, the throughput T is the lowest in such cases. Therefore, there is a trade-off between the throughput and spectrum saving when p varies. The spectrum savings also depend on T_{BS} and, therefore, on the variance of p2p throughputs $T_{p2p,i}$ among subscribers. The largest savings are achieved for the CRWP scenario, followed by the TLW and UDEL scenarios, respectively.

Among the three mobility scenarios, the UDEL is perhaps the closest to many real-world scenarios of urban pedestrian mobility. It also provides the least optimistic outlooks on the feasibility of the mobility-assisted content distribution. Therefore, in the following sections we focus on the UDEL mobility scenario.

4.2 Impact of User Density and Transmission Range

The densities of subscribers and supporters in an area are critical for the performance of the mobility-assisted service. In Fig. 7, we show how the performance changes when the density of subscribers (hence, the relative popularity of the content feed) varies. The overall density of users is kept constant (100 users per km²) and, therefore, increasing the number of subscribers does not affect the total number of users that forward the contents in the p2p network. It however affects the number of subscribers that receive contents from the base stations and, therefore, the number of injection points for new contents. As expected, the throughput increases with the popularity of the feed. The spectrum ratio, however, is not affected because spectrums required by mobility-assisted and infrastructure-based service increase proportionally with the number of subscribers.

In the following scenario, the density of subscribers is 50 per km², as in Fig. 6, but the overall density of users is increased to 1000 per km². The results are shown in Fig. 8 and compared to the results obtained for 100 users per km² shown in the bottom row of Fig. 6. The throughput in Fig. 8 (left) shows a multifold increase compared to the scenario with 100 users per km². The variance of p2p throughputs $T_{p2p,i}$ among subscribers (not shown) decreases, which improves the spectrum efficiency of mobility-

assisted content distribution—the transmission rate in the cellular network T_{BS} is better utilized when each transmitted data disseminates more evenly. Our results indicate that by using only $\sim 10\%$ of the spectrum required for infrastructure-based service the system may achieve throughputs of up to 20% of nominal SRR rate offered by multi-channel radios. However, the density of 1000 users per km^2 might be hard to achieve, even in densely populated urban areas. In the following scenarios, we show how the performance changes for a wide range of user densities and we evaluate the effect of radio interference.

In Fig. 9 (left), we show how the throughput and spectrum ratio change with the density of users for $p = 10\%$ and a transmission range of 10 meters. The results indicate that the throughput T increases linearly with the user density at first. At higher densities, the throughput achieved with the single-channel SRRs is limited by interference. The spectrum ratio remains approximately constant after an initial decrease. This is because the decrease in the variance of p2p throughputs $T_{p2p,i}$, which affects the spectrum usage of mobility-assisted service, becomes negligible when the user density increases beyond ~ 500 users per km^2 . Therefore, when the density of users reaches certain scenario-specific threshold, no additional spectrum savings can be achieved—the throughput, however, could be further increased depending on how the interference is handled.

Fig. 7. Throughput T (left), average transmission rate from the base stations T_{BS} (middle), and spectrum usage ratio (right) achieved with 100 users per km^2 and $p = 10\%$ for various densities of subscribers. Dashed and solid curves correspond to the quasi-static and fluid regimes, respectively. Users move according to the UDEL mobility model.

Fig. 8. Results with 1000 users per km^2 : throughput T (left), average transmission rate from the base stations T_{BS} (middle), and spectrum usage ratio (right). Dashed and solid curves correspond to the quasi-static and fluid regimes, respectively. Users move according to the UDEL mobility model.

Results achieved with transmission ranges of 20 and 30 meters are shown, respectively, in the middle and right columns of Fig. 9. When SRRs use multiple non-interfering channels, the throughput increases both with the transmission range and with the user density. Since our multi-channel model assumes strictly pairwise contacts, increase in user density results in diminishing gains in throughput. When SRRs use a single channel, there is an optimal user density that maximizes the throughput. It depends on the transmission range: The higher the transmission range is, the lower the optimal user density becomes. When the density increases beyond the optimal, the throughput decreases due to interference. The transmission range does not have a significant effect on the spectrum savings, except at the low user densities. At low densities, an increased transmission range may reduce the variance of p2p throughputs among subscribers and, therefore, it may allow a lower transmission rate to be used at the base stations, which improves the spectrum savings of the mobility-assisted content distribution.

4.3 Impact of Limited Cooperation

So far we considered cases where supporters are willing to forward the contents on behalf of subscribers at any time (unlimited cooperation). The actual level of support that subscribers receive will depend on the number of content feeds in the system and the traffic load on the feeds. Here we consider a worst-case scenario of limited cooperation: We assume that all supporters are subscribed to the same feed and they are busy exchanging their own contents whenever they meet each other. Therefore, the supporters will forward the contents on behalf of subscribers only when there are no other supporters

Fig. 9. Throughput (top) and spectrum ratio (bottom) achieved with $p = 10\%$ and the transmission range of 10 m (left), 20 m (middle), and 30 m (right). Dashed and solid curves correspond to the quasi-static and fluid regimes, respectively. Users move according to the UDEL mobility model.

nearby that they can connect to. The cooperation is mutual: when a subscriber and a supporter meet each other, it is randomly decided who is going to support whom.

Results obtained with 50 subscribers per km^2 and a variable number of supporters (total user density varies from 50 to 2000 users per km^2) are shown in Fig. 10. Our previous results for unlimited cooperation are provided in the same figure for comparison. There is a substantial difference in achieved throughputs. The contribution of supporters is still significant: At the user density where throughput is maximized (~ 1000 users per km^2), supporters contribute $\sim 70\%$ of the achieved throughput. As the user density increases further, the throughput decreases due to interference for single-channel SRRs, or it saturates for multi-channel SRRs. Spectrum savings are somewhat lower compared to the case of unlimited cooperation because of the higher variance of p2p throughputs $T_{p2p,i}$ among subscribers, which requires T_{BS} that is relatively high compared to the throughput T .

During the simulations, we kept track of the fractions of time that users spend transmitting/receiving contents and searching for other users. This is used to estimate the energy consumption in the following section.

5 ENERGY CONSUMPTION: A CASE STUDY

Energy consumption is a major concern when designing systems and applications for battery-powered portable devices. Here we evaluate and compare the energy requirements of the infrastructure-based

Fig. 10. Throughput (left) and spectrum ratio (right) achieved with 50 subscribers per km^2 , $p = 10\%$, and transmission range of 10 m. Results for the single-channel and multi-channel SRRs are in the top and the bottom row of the figure, respectively. Dashed and solid curves correspond to the quasi-static and fluid regimes, respectively. Users move according to the UDEL mobility model.

and mobility-assisted content distribution. The performance measure of interest is the energy consumption ratio, which is the ratio of energies required by, respectively, the mobility-assisted and infrastructure-based services to provide a throughput T . Such ratio can be defined in terms of total or peak energy consumption:

- The total energy consumption ratio is the ratio of energies consumed by, respectively, all subscribers and supporters in the mobility-assisted service and all subscribers in the infrastructure-based service.
- The peak energy consumption ratio is the ratio of the energy consumed by the subscriber that is most likely/frequently selected to receive contents from a base station in the mobility-assisted service (this subscriber will be the first to deplete his battery) and the energy consumed by the same subscriber in the infrastructure-based service.

The energy consumption of wireless communication is highly dependent on radio technologies (cellular and short-range) and hardware solutions employed by mobile terminals. Although we have strived to make this study technology-neutral as much as possible, at this point we have to introduce certain assumptions on the radio technologies used in the system: We consider a hybrid 3G/Bluetooth network for mobility-assisted service and we refer to measurement data on energy consumption of 3G and Bluetooth radios available in literature. This is a case study; a general approach to evaluate the energy consumptions of such a hybrid network is outside the scope of this paper. Our methodology may be applied to other short-range radio technologies.

5.1 Infrastructure-Based Content Distribution

Here we consider the energy consumed at a mobile terminal to receive contents via base stations, as described in Section 2. This energy is proportional to the time that the cellular interface of the terminal is in a high-power state. Therefore, it depends on the characteristics of the traffic load: preferably, data is received in short bursts at a high bit-rate to minimize the energy-per-bit ratio. Since the latency requirements of the service are flexible, data can be cached at base stations until the next burst is scheduled for transmission. The cellular interface may remain in an idle state between bursts, where the energy consumption is negligible. In 3G systems, the energy required to receive a burst of data consists of ramp, transfer, and tail energies [31]. The ramp energy is consumed while the cellular interface moves from the idle to the high-power state. The transfer energy is consumed to receive the radio signal and extract the data. The tail energy is consumed while the interface lingers in the high-power state after the transfer is finished, waiting for an inactivity timer to expire. The inactivity timer is set by the operator to minimize the delay involved with transitions from the idle to the high-power state (e.g. in case of short consecutive bursts). Owing to the delay-tolerance of the service, traffic can be shaped

(burst/pause intervals) in such a way that the ramp energy is negligible compared to the transfer energy and inactivity timer is set to zero. Therefore, in the following, we focus on the transfer energy only.

Assume that a multicast transport channel that occupies bandwidth B is used to deliver contents in a 3G network. An HSDPA channel, for example, occupies $B = 5$ MHz of spectrum [32]. This assumption does not affect the spectrum usage of wireless content distribution: the transport channel can be used to deliver other services when idle for this service. Let η be the maximum spectral efficiency on this channel (in bits per second per Hertz), which is shown in Fig. 1 for 50 subscribers per km² and the cell radius of 1000 meters ($\eta \sim 0.5$ bps/Hz). The peak rate on this channel is ηB . To deliver a nominal throughput $T_n \leq \eta B$, it is sufficient to transmit contents during $T_n/(\eta B)$ percent of the time. Let P_{3G} be the battery power consumption of the high-power state of the 3G interface. This power is typically between 1.3 W and 1.5 W [31], [33]. The interface will be in the high-power state only when data are transmitted on the multicast channel. The average energy consumption of the infrastructure-based service is then given by $E_{IB} = (T_n/(\eta B)) \cdot P_{3G}$. When comparing the energy consumptions of the infrastructure-based and mobility-assisted services, T_n is given by the product of the throughput of the mobility-assisted service, T , and the data rate offered by SRRs. Assuming that the mobile devices use Bluetooth v2.1 EDR (Enhanced Data Rate), this data rate is ~ 3 Mb/s. We do not consider cases when cellular backhaul restricts the throughput of the mobility-assisted service, i.e. when $T_n > \eta B$.

5.2 Mobility-Assisted Podcasting

The energy consumed to distribute the contents of a feed consists of energies spent by a user (subscriber or supporter) to i) receive the contents from a base station (E_{BS}), ii) discover other users (E_{ND}), iii) transmit the contents to other users (E_{TX}), and iv) receive the contents from other users (E_{RX}). These energies are estimated as follows.

i) Energy consumed to receive the contents from a base station (E_{BS}) is the same as in the infrastructure-based service at times when a user is selected to receive the contents, and it is zero otherwise. In the quasi-static regime, the same subset of subscribers is selected at all times, as described in Section 3.2. Therefore, for the selected subscribers $E_{BS} = E_{IB}$, while for those that are not selected $E_{BS} = 0$. In the fluid regime, each user is selected with an equal probability p at any time, hence $E_{BS} = p E_{IB}$. For supporters $E_{BS} = 0$ since they are not subscribed to receive the contents.

ii) Energy consumed to discover other users (E_{ND}) depends on the neighbor discovery delay that can be tolerated. The parameters of the neighbor discovery process in Bluetooth can be tuned to provide slow, low-power discovery or fast, high-power discovery. Our assumption in the previous sections was that the discovery does not incur any delay, which is not feasible in practice. However, discovery

delays below one second are almost negligible compared to the typical contact durations at pedestrian speeds with a transmission radius of 10 meters. An extensive study of Bluetooth neighbor discovery algorithms presented in [34] shows that, for Bluetooth v1.2 devices, the battery power needed for an average discovery delay of 0.5 s is $P_{ND} \sim 70$ mW. Allowing an average delay of 1.5 s reduces the power consumption to $P_{ND} \sim 40$ mW. We expect somewhat lower P_{ND} for devices compatible with more recent versions of the Bluetooth standards including v2.1 EDR. The energy consumption E_{ND} is then given by $E_{ND} = P_{ND} t_{ND}$, where t_{ND} is the fraction of time that subscribers spend searching for contacts. This time is determined by simulations. We assume that subscribers search for contacts whenever they are not connected. The energy E_{ND} consumed by supporters is not taken into account as they would have to search for contacts anyway, in order to distribute their own contents.

iii) Energy consumed to transmit the contents to other users (E_{TX}) depends on the battery power needed for the transmission (P_{TX}) and the fraction of time (t_{TX}) that a user spends transmitting the contents: $E_{TX} = P_{TX} t_{TX}$. The battery power consumed during an active transmission in Bluetooth v2.1 EDR is $P_{TX} \sim 65$ mW [35]. The time t_{TX} is determined by simulations.

iv) Energy consumed to receive the contents from other users (E_{RX}) depends on the battery power needed for the reception (P_{RX}) and the fraction of time (t_{RX}) that a user spends receiving the contents from other users: $E_{RX} = P_{RX} t_{RX}$. The battery power consumed during an active reception in Bluetooth v2.1 EDR is $P_{RX} \sim 48$ mW [35]. The time t_{RX} is determined by simulations.

5.3 Performance Results

Our assumptions on parameters needed to calculate the energy consumptions of the infrastructure-based and mobility-assisted content distribution are summarized in Table 1. Note that the energy-efficiency of mobility-assisted service (in bits per Joule) increases with the intensity of data forwarding because battery powers needed for transmission and reception are comparable to the battery power needed for neighbor discovery, as shown in Table 1. This has the following implication: An energy consumption ratio below 100 % indicates that, *at its throughput* T , the mobility-assisted service is more energy-efficient than the infrastructure-based service—this might not be true when the throughput is underutilized (e.g. for light traffic loads).

The peak and total energy consumption ratios for the unlimited and limited cooperation scenarios discussed in Section 4.3 are shown in Fig. 11. These results are complementary to the results on the throughput and spectrum usage shown in Fig. 10 (multi-channel case). The following can be observed for the total energy consumption ratio: In the quasi-static regime (dashed curves), the mobility-assisted service is less energy-efficient than the infrastructure-based service, regardless of the type of cooperation and the density of supporters (there are 50 subscribers per km² and [0, 1950] supporters per km² in

TABLE 1. PARAMETERS FOR THE ENERGY CONSUMPTION CALCULATION

3G/HSDPA		Bluetooth v2.1 EDR	
B	5 MHz	P_{ND}	50 mW
η	0.5 bps/Hz	P_{TX}	65 mW
P_{3G}	1.5 W	P_{RX}	48 mW

our scenario). To the contrary, in the fluid regime (solid curves), the mobility-assisted service is more energy efficient than the infrastructure-based service for most of the considered user densities. The reason is that the fluid regime increases the throughput compared to the quasi-static regime, without incurring any significant increase in energy consumption.

The dependence of the ratio on the supporter density is determined by the energy efficiency of mobility-assisted content distribution. When the number of supporters increases, both the throughput and the energy consumption increase. If the slope of the throughput increase is higher than the slope of the energy consumption increase, the energy-efficiency of the mobility assisted service improves. Hence, the total energy consumption ratio decreases, which is initially the case for all curves shown in Fig. 11 (left). However, at high user/supporter densities, the ratio starts to increase in the quasi-static regime: The energy-efficiency of the mobility-assisted content distribution deteriorates because the energy consumption of supporters increases faster than the throughput.

The extent of cooperation among users has a significant impact on energy consumption. The results indicate that, at low supporter densities, both limited and unlimited cooperations improve the energy efficiency of mobility-assisted content distribution. As the supporter density increases, the gain in throughput becomes negligible in the case of limited cooperation, which is obvious from Fig. 10 (bottom). The increase in energy consumption of supporters is also negligible and, therefore, the energy consumption ratio for the limited cooperation becomes almost constant at high densities. The unlim-

Fig. 11. Total (left) and peak (right) energy consumption ratios achieved with 50 subscribers per km^2 , $p = 10\%$, and transmission range of 10 m. Dashed and solid curves correspond to the quasi-static and fluid regimes, respectively. Users move according to the UDEL mobility model.

ited cooperation, however, may have a positive or negative impact on the energy-efficiency. In the quasi-static regime, unlimited cooperation reduces the energy-efficiency compared to the limited cooperation. However, in the fluid regime, the unlimited cooperation improves the energy-efficiency. This is due to the return in throughput that the unlimited cooperation provides in the two regimes: The slope of the throughput increase is significantly higher in the fluid regime, as shown in Fig. 10.

The peak energy consumption ratio shown in Fig. 11 (right) pertains to the subscribers that are most frequently selected to receive the contents from the cellular network and, therefore, they bear the biggest burden of data forwarding. A particularly high peak ratio would indicate that these subscribers are points of failure, which is the case at low user densities in the quasi-static regime. The peak ratio exhibits similar trends as the total ratio, with one notable exception: The unlimited cooperation in the quasi-static regime improves the energy-efficiency compared to the limited cooperation. This indicates that the increase in throughput outweighs the increase in energy consumption for the selected subscriber, although that is not the case for the total energy consumption.

In summary, our case study shows that the energy-efficiency of mobility-assisted content distribution depends strongly on the selection of subscribers at base stations, which, in turn, depends on channel statistics. Although the system can be forced into the fluid regime by selecting subscribers at random, regardless of their channel condition, that would cancel out any possible spectrum savings. Also, the study shows that cooperation among users subscribed to different feeds might not necessarily have an adverse effect on the energy-efficiency.

6 CONCLUSION AND OPEN ISSUES

We have considered the mobility of people as a supplementary transport mechanism for distributing delay-tolerant contents wirelessly in cellular networks. The mobility-assisted content distribution is a spectrum-efficient alternative to deploying dense cellular infrastructure. We evaluate the performance of the mobility-assisted service in two limiting regimes: quasi-static and fluid. Our results indicate that the mobility-assisted service uses only a fraction of the spectrum required for the infrastructure-based service. The throughput is, however, limited by the efficiency of p2p content forwarding, which depends on the mobility and the number of users in an area as well as the capacity of the used short-range radios. It significantly improves with the number of users, which is a scaling opposite to that observed in cellular networks. Our case study indicates that, in some scenarios, the mobility-assisted content distribution with Bluetooth v2.1 EDR radios can be more energy-efficient than infrastructure-based podcasting. Further improvements might be achieved with low power Wi-Fi and Bluetooth v3.0 radios.

A number of issues still need to be addressed. We evaluated the impact of different cooperation lev-

els, but incentives for cooperation have not been addressed. This is a complex issue related to the power consumption in mobile devices and its optimization. A large number of incentive mechanisms have been proposed for p2p file sharing and message relaying in MANETs (e.g. tit-for-tat, virtual payments, and reputation based schemes), some of which could be applied to our system. The security aspects of the system also need to be addressed. Since mobility-assisted content distribution is not fully controlled by a central authority (in contrast to the infrastructure-based content distribution) there is a danger that a malicious user could infect/harm peer users without the service provider knowing about it. Finally, the performance of the system would ideally be evaluated using traffic models that reflect the way in which new contents are published and consumed by subscribers. The publication and usage patterns of Internet podcasts, which is a form of delay-tolerant contents suitable for the service described herein, have been studied and characterized in [36]. However, these patterns may change dramatically in case of mobile wireless content distribution.

REFERENCES

- [1] Ó. R. Helgason, E. A. Yavuz, S. T. Kouyoumdjieva, Lj. Pajevic, G. Karlsson, "A Mobile Peer-to-Peer System for Opportunistic Content-Centric Networking," *Proc. MobiHeld 2010, ACM SIGCOMM Workshop*, August 2010.
- [2] P. Eugster, P. Felber, R. Guerraoui, Rachid and A-M Kermarrec, "The many faces of publish/subscribe," *ACM Comput. Surv.*, vol. 35, no. 2, 2003, pp 114–131.
- [3] Delay Tolerant Networking Research Group. [Online]. Available: <http://www.dtnrg.org/wiki>. [Accessed March 30, 2012].
- [4] L. Pelusi, A. Passarella, and M. Conti, "Opportunistic Networking: Data Forwarding in Disconnected Mobile Ad Hoc Networks," *IEEE Communication Magazine*, vol. 44, no. 11, pp. 134–141, 2006.
- [5] E. M. Daly and M. Haahr, "The Challenges of Disconnected Delay-Tolerant MANETs," *Elsevier Ad Hoc Networks Journal*, vol. 8, no. 2, pp. 241–250, 2010.
- [6] Huggle. [Online]. Available: <http://www.cl.cam.ac.uk/research/srg/netos/huggle/>. [Accessed March 30, 2012].
- [7] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Impact of Human Mobility on Opportunistic Forwarding Algorithms," *IEEE Trans. Mobile Computing*, vol. 6, no. 6, pp. 606–620, 2007.
- [8] A. Lindgren, C. Diot, and J. Scott, "Impact of Communication Infrastructure on Forwarding in Pocket Switched Networks," *Proc. ACM Workshop on Challenged Networks (CHANTS)*, Pisa, Italy, Sept. 2006.
- [9] N. Banerjee, M. D. Corner, D. Towsley, and B. Levine, "Relays, Base Stations, and Meshes: Enhancing Mobile Networks with Infrastructure," *Proc. ACM MobiCom*, San Francisco, CA, Sept. 2008.
- [10] P. Hui, A. Lindgren, and J. Crowcroft, "Empirical Evaluation of Hybrid Opportunistic Networks," *Proc. Int. Conf. Comm. Systems and Networks (COMSNETS)*, Bangalore, India, Jan. 2009.

- [11] G. P. Perrucci, F. H. P. Fitzek, Q. Zhang, and M. Katz, "Cooperative Mobile Web Browsing," *EURASIP Journal on Wireless Communications and Networking*, 2009.
- [12] A. Balasubramanian, R. Mahajan, and A. Venkataramani, "Augmenting Mobile 3G using WiFi," *Proc. ACM MobiSys*, San Francisco, USA, 2010.
- [13] K. Lee, J. Lee, Y. Yi, I. Rhee, and S. Chong, "Mobile Data Offloading: How Much Can WiFi Deliver?," *Proc. ACM CoNEXT*, Philadelphia, USA, 2010.
- [14] S. M. Hedetniemi, S. T. Hedetniemi and A. Liestman, "A Survey of Gossiping and Broadcasting in Communication Networks," *Networks*, vol. 18, no. 4, pp. 319–349, 1988.
- [15] A. Vahdat and D. Becker, "Epidemic Routing for Partially Connected Ad Hoc Networks," Tech. rep. CS-200006, Duke University, NC, 2000.
- [16] M. Papadopouli and H. Schulzrinne, "Effects of Power Conservation, Wireless Coverage and Cooperation on Data Dissemination Among Mobile Devices," *Proc. ACM MobiHoc*, Long Beach, CA, Oct. 2001.
- [17] J. Chesterfield and P. Rodriguez, "DeltaCast: Efficient File Reconciliation in Wireless Broadcast Systems," *Proc. ACM Mobisys*, Seattle, WA, June 2005.
- [18] A. El Fawal, J.-Y. Le Boudec, K. Salamatian, "Self-Limiting Epidemic Forwarding," Tech. rep. LCA-REPORT-2006-126, EPFL, Switzerland, 2006.
- [19] S. Jung, U. Lee, A. Chang, D.-K. Cho and M. Gerla, "BlueTorrent: Cooperative Content Sharing for Bluetooth Users," *Proc. IEEE Pervasive Computing and Comm. (PerCom)*, White Plains, NY, March 2007.
- [20] W. Zhao, M. Ammar, E. Zegura, "Multicasting in Delay Tolerant Networks: Semantic Models and Routing Algorithms," *Proc. ACM Workshop on Delay Tolerant Networking*, Philadelphia, PA, Aug. 2005.
- [21] W. H. Yuen, R. D. Yates, and S. C. Mau, "Exploiting Data Diversity and Multiuser Diversity in Noncooperative Mobile Infostation Networks," *Proc. IEEE INFOCOM*, San Francisco, CA, April 2003.
- [22] G. Karlsson, V. Lenders, and M. May, "Delay-Tolerant Broadcasting," *IEEE Trans. Broadcasting*, vol. 53, no. 2, pp. 369–381, 2007.
- [23] V. Lenders, M. May, and G. Karlsson, "Wireless Ad Hoc Podcasting," *Proc IEEE SECON*, San Diego, CA, June. 2007.
- [24] V. Vukadinovic and G. Karlsson, "Spectral Efficiency of Mobility-Assisted Podcasting in Cellular Networks," *Proc. ACM Int. Workshop on Mobile Opportunistic Networking (MobiOpp)*, Scottsdale, AZ, Feb. 2010.
- [25] Y. Liang, A. J. Goldsmith, G. Foschini, R. Valenzuela, and D. Chizhik, "Evolution of Base Stations in Cellular Networks: Denser Deployment Versus Coordination," *Proc. IEEE Int. Conf. Comm. (ICC)*, Beijing, China, 2008.
- [26] M. Hata, "Empirical Formula for Propagation Loss in Land Mobile Radio Services," *IEEE Trans. Veh. Tech.*, vol. 29, no. 3, pp. 317–325, Aug. 1980.
- [27] P. Mogensen et al., "LTE Capacity Compared to the Shannon Bound," *Proc. IEEE Vehicular Tech. Conf. (VTC)*, Dublin, Ireland, April 2007.

- [28] I. Rhee, M. Shin, S. Hong, K. Lee, and S. Chong, "On the Levy-Walk Nature of Human Mobility," *Proc. IEEE INFOCOM*, Phoenix, AZ, April 2008.
- [29] UDel Models. [Online]. Available: <http://udelmodels.eecis.udel.edu>. [Accessed March 30, 2012].
- [30] J. Kim, V. Sridhara, and S. Bohacek, "Realistic Mobility Simulation of Urban Mesh Networks," *Elsevier Ad Hoc Networks*, vol. 7, no. 2, pp. 411–430, March 2009.
- [31] N. Balasubramanian, A. Balasubramanian, and A. Venkataramani, "Energy Consumption in Mobile Phones: A Measurement Study and Implications for Network Applications," *Proc. Internet Measurement Conference (IMC)*, Chicago, IL, Nov. 2009.
- [32] 3GPP, "Technical specification group radio access network; feasibility study for OFDM for UTRAN enhancement," Technical Report TR 25.892, V6.0.0, June 2004.
- [33] H. Park, J. Jee, and C. Park, "Power Management of Multi-Radio Mobile Nodes Using HSDPA Interface Sensitive APs," *Proc. Int. Conf. Advanced Comm. Tech.*, Gangwon-Do, South Korea, Feb. 2009.
- [34] C. Drula, C. Amza, F. Rousseau, and A. Duda, "Adaptive Energy Conserving Algorithms for Neighbor Discovery in Opportunistic Bluetooth Networks," *IEEE Jour. Sel. Areas Commun. (JSAC)*, vol. 25, no. 1, pp. 96–107, Jan. 2007.
- [35] B. McFarland (Atheros), "Power-Reduction Details Determine a Chip's Overall Power Consumption," *Wireless Net DesignLine* [Online]. Available: <http://www.wirelessnetdesignline.com/howto/208401866>. [Accessed March. 30, 2012].
- [36] D. Gunawardena, T. Karagiannis, A. Proutiere, and M. Vojnovic, "Characterizing Podcast Services: Publishing, Usage, and Dissemination," *Proc. ACM Internet Measurement Conference (IMC)*, Chicago, IL, Nov. 2009.

APPENDIX

Here we describe steps taken to determine the probing rate T_{PB} and the transmission rate T_{BS} in the quasi-static and fluid regimes

Quasi-Static Regime

In the quasi-static regime there is a distinction between the subscribers that receive contents from the cellular network and the subscribers that must rely on p2p contacts in order to obtain the contents. As discussed in Section 3.2, the throughput of the mobility-assisted service T is given by $T = \overline{T_{p2p,i}}$ for $i \in S \setminus S'$. The transmission rate from the base stations (T_{BS}) should be large enough so that p2p contacts can be fully utilized for the forwarding of the contents, but not larger. It is waste of spectrum to use T_{BS} significantly larger than T because many content items would not be delivered to subscribers that rely on p2p contacts. In an ideal case, $T_{p2p,i}$ of all subscribers $i \in S \setminus S'$ would be equal. Then the system could be engineered so that $T_{BS} = T = \overline{T_{p2p,i}}$ for $i \in S \setminus S'$. However, $T_{p2p,i}$ may vary significantly among the subscribers. T_{BS} has thus to be larger than $\overline{T_{p2p,i}}$ not to limit the subscribers

$i \in S \setminus S'$ that enjoy above-average p2p throughputs, because that would in turn affect the p2p throughputs of other subscribers. Therefore, in the quasi-static regime $T = \overline{T_{p2p,i}}$ and $T_{BS} > \overline{T_{p2p,i}}$ for $i \in S \setminus S'$. The $\overline{T_{p2p,i}}$ depends on the mobility and density of users, the density of subscribers, and the percentage of selected subscribers (p). To select T_{BS} we use the following steps (all data rates and throughputs are from now on assumed to be normalized by the data rate of the SRR):

- i) The selected p -percent of subscribers receive contents at the *probing rate* $T_{PB} = 1$, which is used to gauge the throughput T . The probing rate does not limit the throughput because $T = \overline{T_{p2p,i}}$ and $\overline{T_{p2p,i}} \leq 1$ for $i \in S \setminus S'$.
- ii) We simulate the mobility of users according to the three mobility models described in Section 3.1. We vary the percentage p , the density of subscribers, and the overall density of users. The $\overline{T_{p2p,i}}$ for $i \in S \setminus S'$ is measured for the two physical layer models described earlier.
- iii) We assume that the average transmission rate from the base stations should be equal to the 95-percentile of the p2p throughputs observed among the non-selected subscribers. Hence, $T_{BS} = P_{95}(T_{p2p,i})$ for $i \in S \setminus S'$.

Fluid Regime

In the fluid regime, each subscriber is occasionally scheduled to receive contents from a base station. Missing contents are obtained through p2p contacts with other users. Counted on a unit time interval, subscribers may expect to receive $p \cdot T_{BS}$ data units from base stations and $\overline{T_{p2p,i}}$ data units through p2p contacts. In an ideal case, the average p2p throughput $\overline{T_{p2p,i}}$ would accommodate the missing $(1-p)T_{BS}$ data units, hence $T_{BS} = \overline{T_{p2p,i}} / (1-p) = T$ for $i \in S$. Following the same reasoning as in the quasi-static regime, we conclude however that in practical cases $T = \overline{T_{p2p,i}} / (1-p)$ and $T_{BS} > \overline{T_{p2p,i}} / (1-p)$. To select T_{BS} we follow similar steps as in the quasi-static regime:

- i) The selected p -percent of subscribers receive contents at the *probing rate* $T_{PB} = 1/(1-p)$. The probing rate does not limit the throughput of the system since $T = \overline{T_{p2p,i}} / (1-p)$ and $\overline{T_{p2p,i}} \leq 1$ for $i \in S$.
- ii) *is same as in the quasi-static regime.*
- iii) We select the transmission rate $T_{BS} = P_{95}(\overline{T_{p2p,i}} / (1-p))$ for $i \in S$.
- iv) *is same as in the quasi-static regime.*