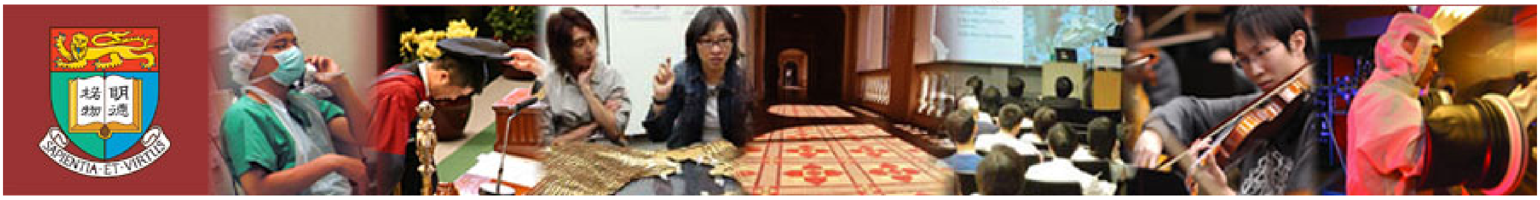

Title	**Cellular traffic offloading through WiFi networks**
Author(s)	**Dimatteo, S; Hui, P; Han, B; Li, VOK**
Citation	**The IEEE 8th International Conference on Mobile Adhoc and Sensor Systems (MASS 2011), Valencia, Spain, 17-22 October 2011. In Proceedings of the 8th MASS, 2011, p. 192-201**
Issued Date	**2011**
URL	**http://hdl.handle.net/10722/158754**
Rights	**IEEE International Conference on Mobile Adhoc and Sensor Systems Proceedings. Copyright © IEEE Computer Society.**

2011 Eighth IEEE International Conference on Mobile Ad-Hoc and Sensor Systems

Cellular Traffic Offloading through WiFi Networks

Savio Dimatteo*†, Pan Hui†, Bo Han†‡, and Victor O.K. Li§
*School of Computer Science, University of St Andrews, St. Andrews, Fife KY16 9SX United Kingdom
†Deutsche Telekom Laboratories, Ernst-Reuter-Platz 7, 10587 Berlin, Germany
‡Department of Computer Science, University of Maryland, College Park, MD 20742, USA
§ Department of Electrical and Electronic Engineering, University of Hong Kong, Pokfulam Rd, HK

***Abstract*—Cellular networks are currently facing the challenges of mobile data explosion. High-end mobile phones and laptops double their mobile data traffic every year and this trend is expected to continue given the rapid development of mobile social applications. It is imperative that novel architectures be developed to handle such voluminous mobile data. In this paper, we propose and evaluate an integrated architecture exploiting the opportunistic networking paradigm to migrate data traffic from cellular networks to metropolitan WiFi access points (APs). To quantify the benefits of deploying such an architecture, we consider the case of bulk file transfer and video streaming over 3G networks and simulate data delivery using real mobility data set of 500 taxis in an urban area. We are the first to quantitatively evaluate the gains of city-wide WiFi offloading using large scale real traces. Our results give the numbers of APs needed for different requirements of quality of service for data delivery in large metropolitan area. We show that even with a sparse WiFi network the delivery performance can be significantly improved. This effort serves as an important feasibility study and provides guidelines for operators to evaluate the possibility and cost of this solution.**

***Keywords*-Cellular traffic offloading, delay tolerant, WiFi access points, trace-driven simulation.**

I. INTRODUCTION

According to Cisco forecasts [1] and practical experiences of mobile operators, we are now facing the “mobile data apocalypse”. Mobile data traffic grows at a compound annual growth rate (CAGR) of 131 percent between 2008 and 2013, and will exceed two exabyte per month in 2013. At the same time, cellular operators in Europe are investing a large amount of money to push machine-to-machine (M2M) communications for billions of machines and smart devices (e.g., automobile and sensors), which will create additional mobile traffic. However, currently cellular networks do not have enough capacity to accommodate such an exponential growth of data. Thus, there is urgency for the research community to look for new solutions.

Operators are rolling out increased bandwidth via High Speed Packet Access (HSPA), Long Term Evolution (LTE) and other upgrades. But simply increasing the speed may not always be economically effective, and there may not be enough bandwidth even with 4G. Moreover, there is always a need to balance end-user satisfaction, infrastructure investments (CAPEX) and operating expenses (OPEX). Even without the mobile data apocalypse issue, if we consider the current flat-rate charging model, cellular operators can still integrate low-cost technologies to reduce the OPEX. Since users are paying a flat rate for the data services, the operators will not gain more from extra consumption of data by the users from their networks. Some operators have realized this issue, and have applied Delay Tolerant Networking (DTN) technologies to transfer bulk data across the Internet [2].

Even in situations where cellular data access is relatively cheap and reliable, the bandwidth offered by such services tend to be orders of magnitude smaller than what is offered by local wireless communication technologies. Thus, it is not appealing to use cellular data networks to transfer large files and operators should balance their traffic between cellular (licensed) spectrum and open (unlicensed) wireless technologies. For example, a better approach to handle this is using the cellular network to transmit the request for some content, and then using delay tolerant techniques to actually deliver the data to mobile devices (this might either be through other mobile nodes as relays, or by simply predicting when a user will come into contact with a high-bandwidth access point and delay the data transmission to that point). The result is that there will be less traffic in cellular networks, which can benefit the operator considerably. At the same time, the local network might be able to provide a content feed with higher quality and lower latency, thus rendering the system beneficial for end-users as well.

In this paper, we propose a DTN approach [3] by leveraging the fact that a significant amount of mobile data are indeed delay tolerant in nature. The target data types are bulk data and videos, which will account for 64% of the world’s mobile traffic by 2013 [1]. Bulk data, for example, large AVI files and data generated by scientific experiments, can usually tolerate certain delay. Certain uplink data created by sensors, and M2M applications such as remote sensing do not require real-time data transmission. By exploiting this intrinsic feature of data, we propose an integrated architecture, Metropolitan Advanced Delivery Network (MADNet), that consists of cellular networks, WiFi networks, and mobile-to-mobile Pocket Switched Networks (PSNs) [4]. We believe that this architecture can provide a low-cost solution in parallel with other solutions like HSPA and femtocells [5].

We focus on metropolitan areas since they have high population density, and high content demands. We consider

978-0-7695-4469-4/11 $26.00 © 2011 IEEE
DOI 10.1109/MASS.2011.26

a scenario in which there is abundant coverage of 3G network in large areas with a flat-rate payment plan for data services. Users have mobile devices that can download and produce rich multimedia contents. The devices have large, but not unlimited, amount of persistent storage. Users may want to upload their content to remote servers. There will be ubiquitous availability of low cost cloud computing resources, and users will upload their files and download media files from the cloud easily. This will be a common scenario in the near future.

For realistic evaluation, we use a real data set of 500 taxis moving in the city of San Francisco for 30 days, and information gathered by the crawling of Youtube videos. We show that with the addition of a limited number of APs, we offload more than 50% 3G cellular traffic and reduce the delay of transfers by more than 50% in the majority of the requests. We further validate the results using another data set collected at completely different time and with different participants. The contribution of this paper is three-fold.

- We propose a novel architecture to leverage opportunistic WiFi and peer-to-peer connections for cellular data offloading. We define the delivery methods for both downstream and upstream, and highlight the application scenarios for applying this new architecture. The architecture is simple, uses only available techniques from current mobile computing research, and hence can be easily prototyped.
- We identify the delay-tolerant natures of certain bulk contents and design the mechanisms to intentionally delay them to achieve the effect of migrating bulk traffic from the cellular network. We confirm the results with two large scale real mobility datasets. We are the first to use such large scale datasets to evaluate metropolitan cellular offloading.
- We quantify the number of APs required for a city-wide WiFi offloading with different quality of service for data delivery. It is obvious that using both cellular and WiFi together will reduce the traffic on the cellular network. The key issue here is how many APs are required to obtain a certain performance improvement. This determines the installation and maintenance costs of the WiFi network, and hence the feasibility of this solution for the operators. In term of research, this strongly motivates the optimization of problems at different levels of detail, such as optimized WiFi deployment, smart caching, multiple concurrent WiFi access, and network management.

This paper does not provide a mathematical formulation of the problem and does not provide prototyping results. Instead, it provides an evaluation of evaluate the feasibility of such offloading solution through real mobility traces, and triggers more specific research topics as the next step. If the ratio of performance improvement and deployment cost is very low, it might not be worth spending more effort on any minor optimizations. Hence, this study is very important for further research on this topic.

The rest of this paper is structured as follows. In Section II, we introduce the general offloading strategies. In Section III, we present MADNet architecture and possible application scenarios. In Section IV, we detail our simulation setup. In Section V, we discuss the simulation results. In Section VI, we review related work. We conclude in Section VII.

II. CELLULAR TRAFFIC OFFLOADING

In this section, we discuss the general solutions for cellular traffic offloading, which include femtocells for indoor offloading, and WiFi and peer-to-peer opportunistic offloading for outdoor and mobile environment. MADNet provides an integrated solution for the latter two cases and uses the cellular network as signaling channel for controlling deliveries.

A. Femtocells for Indoor Offloading

Femtocell technique was initially proposed to improve indoor voice and data services of cellular networks [5]. Femtocells operate on the same licensed spectrum as the macrocells of cellular networks and thus do not require special hardware support on mobile phones. Cellular operators can reduce the traffic on their core networks when indoor users switch from macrocells to femtocells. The disadvantages include the need to install short-range base stations in residential or small-business environments, and the solution is usually for indoor environments and cannot handle macroscopic mobility.

B. Opportunistic Peer-to-Peer Offloading

Han *et al.* [6] proposed to offload traffic from the cellular network to opportunistic peer-to-peer mobile network by selecting k users as the initial set to push the contents. Afterward, the initial set of users aids the propagation of the contents to further users through short-range wireless connectivities (e.g., Bluetooth and ad hoc WiFi). To improve the delivery efficiency, the system can identify the social networks of the users and deliver specific contents to a particular social group [7]. Han *et al.*have shown that even through the simple heuristic of selecting the initial set based on past history, a large fraction of data can be offloaded from the cellular network.

C. WiFi for Outdoor Offloading

WiFi networks operate on the unlicensed frequency bands and cause no interference with 3G cellular networks. WiFi is usually ubiquitously available in urban areas, either deployed by operators as commercial hotspots, shared out as community network (e.g., FON), or deployed by users for residential usage. Meanwhile, there are already several offloading solutions and applications proposed from the industry. For example, the Line2 iPhone application can initiate voice calls over WiFi networks. Recently, Balasubramanian *et al.* [8] proposed a scheme called Wiffler to augment mobile 3G

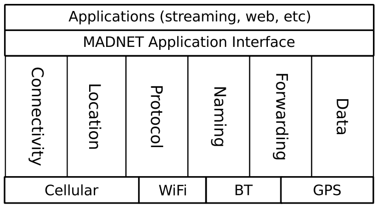

Figure 1. MADNet Architecture

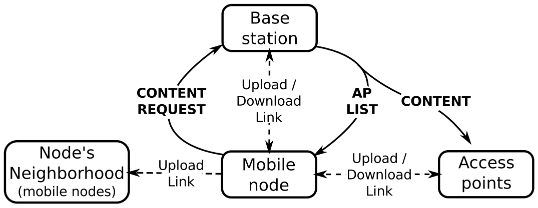

Figure 2. General communication flow in MADNet architecture

using WiFi for delay-tolerant applications. Our focus here is on evaluating the potential costs and gains for providing WiFi offloading in metropolitan area by using large scale real mobility traces for empirical emulation. Lee *et al.*showed promising results of using WiFi for cellular traffic offloading with empirical pedestrian traces [9]. In this paper, we focus on the performance of WiFi offloading on environments with high mobility.

III. MADNET ARCHITECTURE AND SCENARIOS

A. MADNet Architecture

The MADNet architecture is built around the concept of using cellular networks to do signaling and a combination of cellular networks and other communication opportunities to deliver the data. It also uses location services (e.g., GPS, cellular tower triangulation, and WiFi beacon footprints) to help users choose the location of delivery. Figure 1 shows the architecture of MADNet.

MADNet is designed as a middleware between the applications and the physical connectivities. It basically consists of six modules. The *Connectivity* module is responsible for choosing the underlying type of connection for the application. By default for streaming and downloading, it will first choose the cellular networks (we include femtocells in cellular networks), but at the same time it will also launch a local search for available WiFi APs or content. The *Location* module provides location services to applications through the information supplied by GPS, cellular tower triangulation, and WiFi beacon. Users can choose the locations in which they want to pick up the data using a Google map like interface. The *Protocol* module handles the application layer protocols (e.g., HTTP) for data transfer. The *Naming* and *Forwarding* modules are for peer relay purpose. They are responsible for the naming of mobile devices and content, and also the data forwarding for a certain name. The *Data* module is responsible for the assembly of data from different connections (e.g., real time streaming from the cellular network and supplementary data from local search) and pass it to applications.

For local area connectivities, WiFi currently dominates the market for wireless access at homes, offices, and public hotspots for nomadic data services. Provided that pricing levels are reasonable, cellular systems could be competitive also where WiFi networks are easily deployable or existing. Thus, on the one hand, cellular providers want to increase their revenue and are fighting for a bigger slice in the data services segment. On the other hand, they want to reduce operational and network dimensioning costs. To this end, they have to come up with smart ways of handling traffic.

MADNet allows the components to communicate as shown in Figure 2. The system consists of mobile devices that can be carried by people (i.e., nodes). Each node is capable of generating content to upload and requesting content to download. With “content” we refer to a file or a piece of information of interest for the node. When a node wants to download content, it issues a request to the base station currently in range, which replies in two non-exclusive ways: by forwarding the requested content directly through the 3G network, and by using the APs deployed in the city.

The forwarding of the content through the APs may happen in different ways. In this paper, we assume that the requesting node provides some status information about itself (e.g., position, speed, direction) to the base station, which is also aware of the location of the APs (since MADNet is a cellular operator assisted solution). Using this information, the base station predicts the route of the node and produces a list of APs that can serve the node. In parallel, the list is sent back to the node, and the involved APs are fed with the content. Finally, the node consults the list to connect to the serving AP and download the content. The upload procedure works similarly, except that no status information is sent by the node. After downloading the content, the node keeps it in its storage to facilitate future peer-to-peer downloading by its neighbors through local connections such as Bluetooth or WiFi.

B. Advanced Deliveries

In MADNet, the delivery methods can be classified into two categories: downstream and upstream.

1) Signaling and Pickup: This is the basic method for downloading bulk content that is not time-critical. Users will select some content to download and use cellular networks to initiate fetching of this content. The user interface will also inform users of WiFi APs in the user’s vicinity. Users can choose one or more APs where they want the data to be available for pickup, based on their itinerary (Figure 3). The system will then move the data to those particular APs using the backbone network, and the user can pick it up when she arrives at the AP using WiFi.

We further divide this category in four sub-categories:
(a) *Complete Oracle*: This is the upper bound for the performance of opportunistic delivery, in which service providers

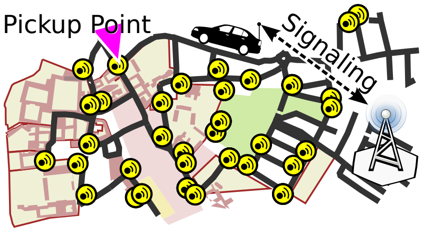

Figure 3. Signaling and pickup scenario

are able to accurately predict user movement and can ship the content to the hotspot to be visited in advance. This may be practical if MADNet uses cellular networks as a control channel to instruct servers where to deliver the data.

(b) *Regular Oracle*: In this case, the system cannot predict the complete movement of users but it knows their regular traveling patterns and can predict the probability of their approximate positions. This is practical if a central server can learn the regularity (if any) from user mobility history.

(c) *Popular Hotspots*: Some popular content is shipped and cached at hotspots in advance, or the content is shipped to several popular hotspots when the content is requested by a certain user. In this case, we assume the central server does not have any knowledge about the location of users.

(d) *Pure Opportunistic*: We assume a certain distribution of media files, for example, on other mobile devices, and when a user requests some content from the network, the MADNet software on the mobile devices will also issue a local search.

2) Peer Relay: The upstream category can be divided into two sub-cases. In the first case, the entities that are uploading a file are mobile. The files/data to be uploaded may be user-generated contents, for example, pictures and videos produced by the handhelds. The alternative way of uploading data for this case is that users can always wait until they arrive to places with reliable wireless connections (e.g., home or office). The only issue for this case is the incentives to use WiFi instead of persistent 3G uploading. One possible reason that discourages a user from using cellular networks for this purpose is battery consumption. Uploading several data on the go through cellular networks drains the battery of mobile phones very quickly. Operators can provide good femtocell coverage in residential areas so users can wait until they arrive to these comfort zones, where their phones will have a power supply and can sustain the transmission of large amounts of data. Indeed, the contents can also be carried through store-and-forward by other mobile devices.

In the second case, the users have no or very limited mobility. A typical example is the sensor network for environmental monitoring. In this case, the entities cannot move to places with Internet connectivity so it seems that 3G network would be a good alternative to move the data to centralized servers. The sensors can upload the data to mobile devices which pass by (e.g., normal pedestrians) or to researchers who visit regularly (i.e., data mules [10]). The data can then be passed to the cloud/server by multi-hop transmissions or via WiFi hotspots.

C. Application Scenarios

We consider three application scenarios: high quality video-on-demand, bulk data transfer, persistent uploading.

1) High Quality Video Streaming: We propose to use opportunistic networking to improve the quality of videos streamed to the user. Currently, users connect to well-known online services (e.g., Youtube.com) and download desired video via 3G networks. Instead, MADNet introduces the possibility to search the same video within the users' neighborhood or their local network (i.e., WiFi APs). If the content is available locally, the first chunks of the video can be buffered directly. Otherwise the video will be streamed through cellular networks immediately and subsequent chunks will be shipped to the predicted APs and will be downloaded opportunistically. There is no negative impact on user experience since the streaming from 3G networks proceeds as usual during the local buffering and, at the same time, the local network or users' neighborhood feeds the content with higher quality and lower latency than what is normally possible.

2) File Sharing and Bulk Data Transfer: In this case, we look at bulk data which is not time critical and therefore there is no need to contend for bandwidth with time-critical applications. Through MADNet, mobile data can be downloaded in a fully opportunistic fashion. The proposed delay tolerant techniques include pure opportunistic forwarding between mobile nodes and prefetching the content to those WiFi hotspots the users are predicted to pass in the future. It has been shown that utilizing opportunistic communication can greatly improve performance of content dissemination in networks of sparse infrastructure [11].

3) Persistent Uploading: In a somewhat different application scenario, we consider file uploads instead of downloads. In this scenario, users produce large quantities of content that, on the long run, exceed the storage capacity of their mobile devices. For example, people use their mobile phones to take pictures and record videos, and use 3G networks to upload their content to cloud storage (e.g., Facebook mobile upload). In this case, the wide availability of flat-rate 3G connection allows the users to not wait until they go home and use the cable for the upload; they can share their files anywhere, immediately and automatically. The persistent uploading scenario can also be found in sensor networks used to monitor urban environments. A sensor can take pictures, record videos and store other environmental information during certain intervals of time. There can be many monitoring sites in a city and it would not be cost-effective to install a wired or even a mesh network to connect these sites. With the increase in popularity of M2M applications in the future, a lot more uplink traffic may be generated and potentially loaded into cellular networks.

IV. TRACE-DRIVEN SIMULATION

We focus on the evaluation of opportunistic WiFi offloading in order to offer an in-depth analysis and refer interested readers to Han *et al.* [6] for the evaluation of cellular traffic offloading through opportunistic communications. We evaluate the performance of MADNet through an event-driven simulator. The simulator provides the visualization of the current simulation state into an interactive environment. This environment displays the current position of the users, the APs deployed, and the simulation events on the Google map of San Francisco. Cellular base stations are abstracted as single entities equipped with storage memory and computational capabilities.

A. Data Set

We used a publicly available data set to determine the mobility of the users. The dataset contains the movements of 536 taxis, that were recorded in the city of San Francisco as part of the *Cabspotting* project [12]. For a given vehicle, its GPS location and occupancy were recorded over a period of 30 days. However, these spatio-temporal data were not sampled at regular intervals over time. Therefore, we exclude from the data set all the taxis with more than 100 seconds average sampling frequency and standard deviation $>1,000$. We obtain a set of 343 taxis having a better sampling frequency. We perform our evaluations using this data set because it is less likely to include measurement errors.

B. Download Requests

A *download request* (DR) is a file or content requested by the user for download. Inside the simulator, we represent it as a tuple of parameters ⟨ *id, timestamp, size, lookahead* ⟩. Every time a user requests a certain content he also declares a "final location" and the time at which he will reach that location. We name it as *lookahead time* and define that, once expired, the requested content will no longer be of interest for users. To define the time at which a DR is issued (i.e., timestamp) and the lookahead time, we use the occupancy information inside the traces. In particular, we define that a user issues a DR when the taxicab is first occupied and the lookahead time expires when the vehicle is free. Note that the choice implies no simultaneous DRs per node. Figure 4 shows the cumulative distribution function of lookahead time extracted from the raw data set. As can be seen, the lookahead time exhibits a log-normal distribution having the majority of DR lasting between 100 and $1,000$ seconds. This is reasonable if we think that the lookahead time reflects the occupancy of the taxis in San Francisco.

We say that a DR is "satisfied" when the requesting node has received all the bytes of requested content. We call B_r the size in bytes of every requested file. As we increase B_r, we obtain longer satisfaction time or no satisfaction at all. To make the choice, we run one simulation within 3G networks using the same general settings of our final simulations (see Section IV). In particular, we evaluate the behavior of 3G networks when all the requests are not satisfied (their lookahead always expires or, ideally, $B_r{=}\infty$).

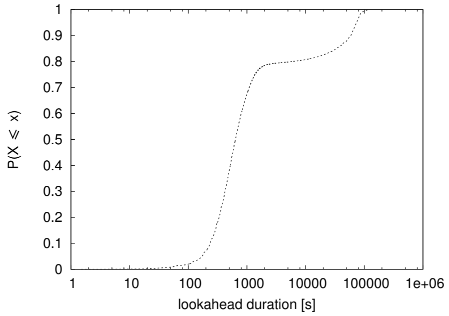

Figure 4. Lookahead time distribution

The downloadable amount of data is log-normally distributed around a mean value of 29.23 MB. By choosing B_r according to such distribution we can determine the behavior of 3G networks also in accordance with the lookahead time. In our study, we model the future situation in which 3G networks can only satisfy a limited number of requests. Therefore, we set $B_r{=}29.93$ MB, that is, only 35.3% of the DRs issued will be satisfied through cellular networks.

C. Upload Requests

We define a *point of interest* (POI) to be an attractive situation/object/place for a user. When a user enters the vicinity of one POI (i.e., within 50 meters in our case), he records and uploads a video through his mobile device. For simplicity, other POIs are ignored by the user when he is uploading the data. We represent an *upload request* (UR) as a pair ⟨*time, size*⟩, where "time" is determined at simulation run time and defines *when* the UR is issued, and "size" is specified for each POI and characterizes the number of bytes of generated content. Also, we say that an UR is "satisfied" when a user uploads all the bytes of generated content.

1) POI Placement: To define the positions of POIs in San Francisco we implement a "video information grabber" program based on Youtube Data APIs. Such APIs offer the possibility to search for videos uploaded and geo-tagged (i.e., the video is associated with a particular location on the earth) by Youtube users. The search operation is performed by passing a central point P and a radius (e.g., $10km$) as parameters to the API. The API returns a list of information about videos that are geo-tagged around P and lie within the requested range. We next configure our program to scan for videos within a rectangular area around San Francisco. In particular, we top-down scan a $554km^2$ large area from west to east. At each step of the scanning procedure, we shift P by 48 meters and search for videos geo-tagged within 500 meters from P. In the scanning procedure, we use small radius values and include overlapping areas because the Youtube API returns a maximum of 50 results. Therefore, when the scan finishes, we remove duplicated coordinates and obtain a total of $8,315$ geographically-placed POIs.

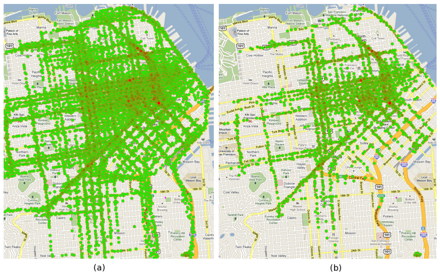
(a) (b)

Figure 5. Deployment of access points in the simulator, with thresholds $1,000$ (a) and $3,000$ (b)

2) Size: In addition to the GPS coordinates of the videos, during the scanning procedure (see Section IV-C1) we also grab their duration. We assume that all the videos are recorded through a conventional mobile phone (e.g., Nokia N73), and calculate their size according to the associated video encoding parameters. In our case, we suppose that the videos are encoded with a resolution of 352×288 pixels in MPEG-4 video and AAC audio (1 channel). With these settings, we obtain video sizes ranging from 103 KB to a maximum of 743.53 MB, with an average of 19.43 MB. Note that in this case we do not grab directly the size of the videos, and assume that all of them were recorded through a mobile phone.

D. Access Point Deployment

In line with known findings about quality of service of the Internet [13], [14], we say that a node is *happy* if it downloads at least one byte of the desired information within 10 seconds.[1] Initially, we run one simulation without deploying any APs in San Francisco. In this run, the nodes request the content but the content is never delivered to them. When a node becomes unhappy we add a new AP into a ranked list, set its location to the location of current node , initialize its rank to 1 and change the node status to “happy”. We repeat the procedure each time a node becomes unhappy, but before adding an entry to the list, we check if the current position of the node falls within the range of existing APs. In other words, new APs are added to the list only when none of the previously stored APs has the node in its range. The rank of the matching AP in the list is incremented by one. At the end of the simulation we dump the resulting list (i.e., a list of ranks and positions) to the deployment file used as input in future simulations.

In the above described procedure we increase the rank of an AP each time it is needed by a node, thus the final array of ranks is representative of the AP usage. Deployment is performed according to a “threshold” (T), that is, a reference value used to consider APs with greater or equal rank in the simulation. We observe that T is roughly in inverse proportion with the centrality [15] of the deployed APs: as T increases, less APs are deployed around the main streets of San Francisco. Instead, when T decreases, more APs are gradually deployed from central to marginal streets.

Table I

THRESHOLDS AND CORRESPONDING NUMBER OF DEPLOYED APS

Threshold	**# of Deployed APs**
$1,500$	$4,271$
$3,000$	$1,948$
$10,000$	178
$12,000$	82
$13,000$	54
$14,000$	41

For example, by setting T=$1,000$ we deploy $6,106$ APs as shown in Figure 5a, while by increasing the threshold to $3,000$ we deploy $1,948$ APs as shown in Figure 5b. In Table I, we summarize the thresholds we set in the simulations and the corresponding number of deployed APs.

E. 3G and WiFi Infrastructure

Empirical studies on 3G network capacity and performance assert that it is practically impossible to predict the actual capacity of 3G cells based on known theoretical models and standard parameters [16]. Despite the level of complexity that can be reached in simulating 3G cells, we decide to adopt a simple model in which there is one “ideal” cell that is permanently connected to each moving node in the simulation (i.e., infinite coverage). This cell is always able to supply the nodes with a constant download and upload bandwidth of 370 Kbps and 64 Kbps, respectively. We choose these values based on our measurement of T-Mobile’s UMTS 3G network in Berlin, Germany. The motivation for this choice goes beyond keeping the simulator computations low in complexity. This not only reproduces the 3G full-coverage assumption described in Section I, but also allows us to present our results when the 3G network performs globally at its best.

In order to analyze our system in the worst case, we want the APs to be more realistic than the ideal 3G cell. To do this, we use results from empirical measurements of the WiFi capacity [17]. The authors evaluate the fluctuation of the upload/download throughput perceived by the user in several urban environments. Several measurements were taken outdoor in the range of $5-75m$ from the indoor-deployed AP. To adapt such results to our simulator, we assign the throughput to every user below a distance of $5m$: as a node gets closer to an AP, it can transfer up to the maximum value (B_{max}) of 18 Mbps. For each AP we also define a policy to manage the bandwidth across the node requests. A node that requires a bandwidth of B bps can obtain it only if there is enough bandwidth available. We define such quantity as:

$$B_{available} = B_{max} - \sum_{i \, \in \, D} B_i \ ,$$

where D is the set of the nodes connected to the AP and B_i is the bandwidth reserved for the i-th node of the set. When

---

[1] Indeed this is a lower bound for the network performance as the speed of the current wireless network has notably increased since the publication of our cited literatures

Table II

SUMMARY OF SIMULATION CASES AND CORRESPONDING SETTINGS

Case	Run ID	Network	Threshold (T)
Download (BDT and VST)	D0	3G	∞
Download (Bulk)	D1	WiFi	$1,500$
	D2	WiFi	$3,000$
	D3	WiFi	$10,000$
	D4	WiFi	$12,000$
	D5	WiFi	$13,000$
	D6	WiFi	$14,000$
Download (Video Streaming)	D7	WiFi + 3G	$1,500$
	D8	WiFi + 3G	$3,000$
	D9	WiFi + 3G	$10,000$
	D10	WiFi + 3G	$12,000$
	D11	WiFi + 3G	$13,000$
	D12	WiFi + 3G	$14,000$
Upload (MT1 and MT2)	U0	3G	∞
Upload (Multimedia 1)	U1	WiFi	$1,500$
	U2	WiFi	$3,000$
	U3	WiFi	$10,000$
	U4	WiFi	$12,000$
	U5	WiFi	$13,000$
	U6	WiFi	$14,000$
Upload (Multimedia 2)	U7	WiFi + 3G	$1,500$
	U8	WiFi + 3G	$3,000$
	U9	WiFi + 3G	$10,000$
	U10	WiFi + 3G	$12,000$
	U11	WiFi + 3G	$13,000$
	U12	WiFi + 3G	$14,000$

the available bandwidth is less then B but still greater than zero, it is entirely assigned to the requesting node. The AP refuses the connection of a node each time $B_{available} = 0$.

F. Setup Plan

The 26 settings used in our simulations are summarized in Table II. The download case can be divided in two sub-cases: video streaming transfers (VST) and bulk data transfers (BDT). The upload case can also be divided in two sub-cases: "Multimedia 1" (MT1) and "Multimedia 2" (MT2). In MT1 the users upload the content through the WiFi network only. In MT2, the content can be uploaded through WiFi and the 3G network. In both MT1 and MT2, the users generate the same content. A base station receives an (upload or download) request after 3 seconds since it is issued by a node. After the request is received, the list of APs that can serve the request is sent to the node in 2 seconds. The base station can deliver the content to the APs in 1 second simultaneously. We have chosen these timings on an arbitrary basis, and prefer to leave the investigation of the exact engineering parameters for future work. All the download requests issued by the nodes have a fixed size of 29.93 MB. According to the occupancy information in the traces, The nodes begin to issue the download requests on 17 May 2008 at midnight. After this date, the number of download requests grows linearly over time up to $331,075$, with an average rate of one request every 6.28 seconds. The number of upload requests is decided at run time during each simulation run. Because the same traces are used, the nodes try to download the same content in each download case.

V. EVALUATION RESULTS

In this section, we compare the performance of MADNet against 3G networks when users download and upload data. We report the results of the two cases separately, characterizing the system, considering the satisfaction of users, delays and network load.

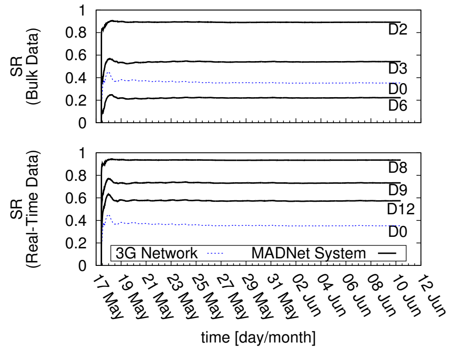

Figure 6. Variation of the download satisfaction ratio over time with different settings

A. MADNet Download

1) Satisfaction of the Users: We define the *satisfaction ratio* (SR) to be $\frac{R_s(t)}{R_i(t)}$, where $R_s(t)$ is the total number of satisfied (download or upload) requests at time t, and $R_i(t)$ is the total number of requests issued by users at time t. We evaluate SR over time in case of download (D0–D12). To enhance readability, in Figure 6 we show only the evaluations for D0, D2, D8, D3, D9, D6, D12. As can be seen, SR remains stable in each simulation run after a period of two days. According to the traces, not all the taxis are moving or requesting data during this period.

In case of video streaming, the requests are always better satisfied than inside the 3G network. Therefore, deploying more access points increases the number of satisfied requests. However, such improvement is not directly proportional to the number of deployed access points. For example, introducing 41 APs into 3G networks (i.e., D12) increases SR by 21.49%, while an improvement of 29.46% is given by adding 82 APs (i.e., D10) in the city. Also, deploying a large number of APs (e.g., D7) we obtain a high increase of the average SR of 60.92%, but more than half of the increase is contributed by 4.15% of the APs. As the threshold also determines the centrality of deployed APs (see Section IV-D), we conclude that a larger number of requests can be satisfied by strategical deployment of APs along the central streets of urban area.

Similar observations can be made for the bulk data transfer case. However, in such case if we deploy too few APs (e.g., D5 and D6), the requests cannot be satisfied with the same ratio as in 3G networks. Despite this, we find that about the same satisfaction ratio of 3G networks can be obtained through 82 APs. With respect to the overall results, we observe that decreasing the threshold keeps the SR at about the same value for both cases. In fact, for a given threshold the presence of cellular networks is the only difference between the two cases. Therefore we conclude that deploying more APs (that are gradually placed in less central streets) makes the SR less influenced by cellular networks.

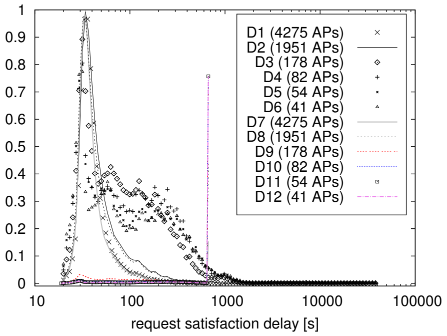

Figure 7. Request satisfaction delay distribution of download requests

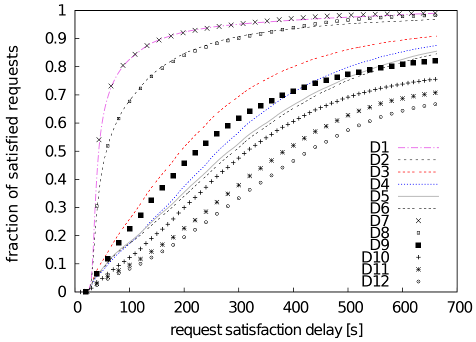

Figure 8. Satisfied requests ratio across different settings

2) Delays: In Figure 7 we show the distribution of the time by which the download requests (of fixed size) are satisfied. In D0, every request is satisfied in 663 seconds as the 3G cells continually supply constant throughput to users. This timing is an upper bound for the requests issued in each video streaming case (i.e., D7–D12) as, in these cases, the cellular network is combined with the WiFi network. In fact, requests that were not satisfied through the WiFi due to a lack of coverage are satisfied through the cellular networks anyway. In the bulk data transfer cases the network consists of APs only and therefore some requests are satisfied after longer time. In Figure 7, this is represented by the "tail" in nearly all the considered cases.

We further examine the effects of the above-described delay before 663 seconds. In Figure 8, we plot the fraction of the requests satisfied during the whole simulation as function of the above-described delay. As can be seen, in MADNet the DRs start to be satisfied after 20 seconds and at least 67% of them are satisfied before 663 seconds. Also, we observe a sudden increase of this performance for low thresholds. For example, more than 80% of the satisfied requests are satisfied by 200 seconds in D1, D2, D7 and D8.

3) Network Load: The total amount of data downloaded in the cellular network grows linearly at an average rate of 28.16 MBps. Given that we obtain a similar growth in each download case, in Table III we report the average gain of this quantity with respect to D0. The negative gain for D3 and D4 is due to few APs deployed in the city (e.g., less than 55). In D5 and D6 we obtain a negative gain because a portion of the data transferred in D0 do not contribute to the satisfaction of the requests. In other words, in D0, the lookahead is reached before the transfer has completed. For scenarios in which more than $1,940$ APs are deployed (i.e., D1, D2, D7, D8), we obtain more than 25% gain over D0. Of course, we observe the highest gains in cases that involve transfers through the 3G network and the WiFi network. We further characterize the VST case by looking at Figure 9, where we show the average load of the 3G and WiFi networks over time for VST and MT2 cases. As can be seen, in the VST case, a small number of APs can contribute positively in offloading the cellular network. For example, by deploying 178 APs (i.e., D9) the ideal 3G network is offloaded of about 50% of the traffic. In case of BDT, the general considerations for the SR variation apply.

Table III
AVERAGE GAIN OF DOWNLOADED DATA OVER THE CELLULAR NETWORK

Run ID	Avg Gain over D0
D1	31.55 %
D2	25.74 %
D3	−15.52 %
D4	−33.28 %
D5	−45.34 %
D6	−52.44 %
D7	34.72 %
D8	32.49 %
D9	20.2 %
D10	15.9 %
D11	13.38 %
D12	11.53 %

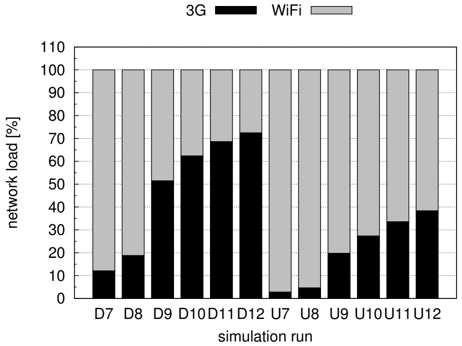

Figure 9. Average load of 3G and WiFi networks for VST and MT2

B. MADNet Upload

1) Satisfaction of the Users: We obtain SR=1 in each upload simulation run. Indeed, this does not indicate that the users' requests are equally satisfied in each scenario. U0 here makes no exception. Therefore, we evaluate the average gain in number of satisfied requests with respect to U0 and summarize the results in Table IV. As can be seen, the number of additional satisfied requests may grow up of one order of magnitude. We obtain a noticeable growth of this quantity whereas more than $1,948$ APs are deployed, suggesting that a significant increase of satisfied requests can be obtained by deploying few central APs.

2) Delays: We now examine the time required to satisfy the URs with respect to the 3G network. Figure 10 shows

Table IV
AVERAGE GAIN OF SATISFIED UPLOAD REQUESTS ON THE CELLULAR NETWORK

Run ID	Avg. Gain of upload $R_s(t)$ over U0
U1	1,162%
U2	917.58%
U3	327.17%
U4	226.73%
U5	187.16%
U6	157.21%
U7	1,183.26%
U8	949.7%
U9	391.98%
U10	300.16%
U11	264.07%
U12	241.03%

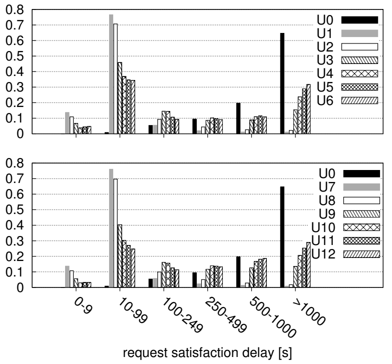

Figure 10. Request satisfaction delay distribution of upload requests

how such delays are distributed within some intervals of time. The y-axis shows the fraction of requests that are satisfied in the corresponding interval of time reported on the x-axis. We can see that, in the cellular network (U0), 85% of the URs are satisfied after about 8 minutes and 19 seconds. Within U1, U2, U7 and U8, this delay is reduced by 80% for the majority ($>80\%$) of the URs. We highlight that the request satisfaction delay in MT2 cannot be worse than U0 because in the worst case the URs are satisfied with the same timings of U0.

3) Network Load: The MT2 case can be further detailed. As can be seen in Figure 9, the results are similar to the download case: few APs can contribute significantly in off-loading the cellular network. We also observe that the 3G network is here better off-loaded that in the BDT case. In this case, as the throughput offered in U0 is limited for each user (64 Kbps), the upload requests are soon satisfied by the WiFi network.

C. Scalability

We have evaluated the average percentage of data downloaded or uploaded through WiFi and 3G networks as we gradually introduce the 500 taxis into subsequent simulation scenarios. Each of these scenarios consists of $4,271$ APs, and is set up as outlined in Section IV-F. According to the results, increasing the number of taxis loads the WiFi networks and offloads the 3G networks of about 0.5% of the downloaded traffic. We obtain the opposite behavior for the upload case. However, this slight increase indicates a good degree of stability of the system in a given scenario. We are currently working on a theoretical framework for the scalability issues when the system scales up to many more users.

D. Confirmation of Results

We verify our results using a new data set collected through the Cabspotting API [12]. The dataset contains the movements and the occupancy of 387 taxis that operated in San Francisco in January 2010. The average sampling frequency of the positional observations of a taxi ranges from 100 to 110 seconds, with a standard deviation between $8,000$ and $10,000$ seconds. We run the simulations with the same settings and parameters described in Section IV-F, and obtain slightly better results than in our previous evaluation. This suggests that the principles used to place the APs and the POIs remain valid even when the movement and patterns of content request of the taxis varies.

VI. RELATED WORK

Goemans *et al.* [18] proposed an architectural and protocol framework that allows 3G providers to offload cellular traffic while distributing the content efficiently. In this work, the offload is performed by caching the content in the infrastructure of "resident subscribers" (i.e., almost static users). Despite our work shares the adoption of a fixed infrastructure, the authors' contribution is focussed on game theoretical aspects of user cooperation rather than on effects resulting from the metropolitan mobility of the nodes. Also, a scenario in which the users can produce and upload contents is not addressed.

The scenario discussed in our paper is similar to Unified Cellular and Ad-Hoc Network (UCAN) [19]. UCAN was designed to increase the throughput of the cell and maintain fairness between users. Similarly to our work, UCAN mobile clients rely on the combination of WiFi and cellular networks. However the paper focusses on improving the throughput of cellular networks by evaluating specific protocols, rather than trying to offload the traffic in a static infrastructure. Wiffler [8] is a system to augment 3G with WiFi for vehicular networks, which is evaluated mainly for a small and sparsely populated city and does not show system scalability when the number of users scales up to several hundreds.

There are also several efforts that deal with specific problems of information dissemination. Aioffi *et al.* [20] proposed a mobile dynamic content distribution network model that aims to reduce the total traffic in enterprise networks. To achieve this, their model takes into account the variation of throughput demands in order to decide whether to remove or keep the replicas in the network. In [21], content dissemination is analyzed as an optimization problem when a small amount of information (e.g., traffic update) is required to be transmitted in a timely manner to a set of nodes passing along the same path. Although we have not considered the exact mechanisms of information dissemination across users, we do not exclude the applicability of such approaches to further reduce the network load. A solution that can probably better fit into our approach is the application of content dissemination strategies to enhance user experience with multimedia services. For example, in [22], servers are able to receive feedback from rich-media applications of the users and adapt the timings and the quantity of data delivery to each handheld.

Overall, it is expensive to deploy a city-wide WiFi network only to offload cellular data. As we mentioned in our previous work, the cellular network operators, WiFi service providers, and end-users (with already deployed residential WiFi) should cooperate to support mobile data offloading, as it is win-win-win for them [23].

VII. CONCLUSIONS

In this paper, we present an architecture for the integration of WiFi networks and mobile-to-mobile Pocket Switched Networks (PSN) with cellular networks to provide a low-cost solution to handle the exponential growth of mobile data traffic. Using real mobility traces from the city of San Francisco, we have shown that only few hundreds of WiFi APs deployed in an area of $313.83km^2$ can offload half of the mobile data from the 3G network in our scenario settings for both download and upload cases. The MADNet architecture is simple, uses commonly available techniques from current mobile computing research, and can be easily prototyped and deployed using off-the-shelf hardware equipment. Although the results are encouraging and suggest the feasibility of opportunistic data offloading, we believe this is still a fundamental step toward full integration of opportunistic networks with cellular networks. Much more research is required for practical use. We will study efficient data replication and caching schemes, which can reduce the delays induced by transferring data to the APs. We plan to examine centrality metrics used to study spatial urban settings and observe whether any effective deployment scheme can be obtained by considering only the topological structure of the city. For participants that share WiFi APs or relay data, we can explore potential incentive schemes among the cellular operators, normal users (residential WiFi owners), and fixed-line operators (hotspot owners). The practical use of the proposed architecture involves additional research and engineering issues, including authentication and accounting, signalling and transportation layer design. We believe that our work will trigger many coming research challenges.

ACKNOWLEDGMENT

Bo Han was supported in part by NSF Award CNS 1010789.

REFERENCES

[1] (2009, January) Cisco visual networking index: Global mobile data traffic forecast update.

[2] N. Laoutaris, G. Smaragdakis, R. Sundaram, and P. Rodriguez, "Delay-Tolerant Bulk Data Transfer on the Internet," in *ACM SIGMETRICS 2009*, June 2009.

[3] K. Fall, "A delay-tolerant network architecture for challenged internets," in *ACM SIGCOMM*, August 2003.

[4] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Impact of human mobility on opportunistic forwarding algorithms," *IEEE Transactions on Mobile Computing*, vol. 6, no. 6, pp. 606–620, 2007.

[5] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: A survey," *IEEE Communications Magazine*, vol. 46, no. 9, pp. 59–67, September 2008.

[6] B. Han, P. Hui, V. Kumar, M. V. Marathe, G. Pei, and A. Srinivasan, "Cellular Traffic Offloading through Opportunistic Communications: A Case Study," in *ACM CHANTS 2010*, Sep. 2010.

[7] P. Hui, J. Crowcroft, and E. Yoneki, "Bubble rap: social-based forwarding in delay tolerant networks," in *Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing*, ser. MobiHoc '08. New York, NY, USA: ACM, 2008, pp. 241–250.

[8] A. Balasubramanian, R. Mahajan, and A. Venkataramani, "Augmenting Mobile 3G Using WiFi," in *ACM MobiSys 2010*, Jun. 2010, pp. 209–222.

[9] K. Lee, I. Rhee, J. Lee, S. Chong, and Y. Yi, "Mobile data offloading: how much can wifi deliver?" in *Proceedings of the 6th International COnference*, ser. Co-NEXT '10. New York, NY, USA: ACM, 2010, pp. 26:1–26:12.

[10] S. J. Rahul C Shah, Sumit Roy and W. Brunette, "Data mules: Modeling a three-tier architecture for sparse sensor network," in *IEEE Workshop on Sensor Network Protocols and Applications (SNPA)*, May 2003.

[11] P. Hui, A. Lindgren, and J. Crowcroft, "Empirical evaluation of hybrid opportunistic networks," in *IEEE COMSNETS 2009*, January 2009.

[12] "Cabspotting," http://www.cabspotting.org. [Online]. Available: http://cabspotting.org

[13] A. Bouch, A. Kuchinsky, and N. Bhatti, "Quality is in the eye of the beholder: meeting users' requirements for internet quality of service," in *ACM CHI '00*, 2000, pp. 297–304.

[14] A. King and J. Nielsen, *Speed Up Your Site: Web Site Optimization*. Pearson Education, 2003.

[15] P. Crucitti, V. Latora, and S. Porta, "Centrality measures in spatial networks of urban streets," *Physical Review E*, vol. 73, p. 036125, 2006.

[16] W. L. Tan, F. Lam, and W. C. Lau, "An empirical study on the capacity and performance of 3g networks," *IEEE Transactions on Mobile Computing*, vol. 7, no. 6, pp. 737–750, 2008.

[17] M. Solarski, P. Vidales, O. Schneider, P. Zerfos, and J. P. Singh, "An experimental evaluation of urban networking using ieee 802.11 technology," in *IEEE OpComm 2006*, September 2006.

[18] M. Goemans, L. E. Li, V. S. Mirrokni, and M. Thottan, "Market sharing games applied to content distribution in ad-hoc networks," in *ACM MobiHoc '04*, May 2004.

[19] H. Luo, X. Meng, R. Ramjee, P. Sinha, and L. E. Li, "The design and evaluation of unified cellular and ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 6, no. 9, pp. 1060–1074, 2007.

[20] W. M. Aioffi, G. R. Mateus, J. M. de Almeida, and A. A. F. Loureiro, "Dynamic content distribution for mobile enterprise networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 10, pp. 2022 – 2031, Oct. 2005.

[21] M. Aezladen, R. Cohen, and D. Raz, "Locally vs. globally optimized flow-based content distribution to mobile nodes," in *IEEE INFOCOM 2009*, April 2009.

[22] R. Bhatia, G. Narlikar, I. Rimac, and A. Beck, "Unap: User-centric network-aware push for mobile content delivery," in *IEEE INFOCOM 2009*, April 2009.

[23] B. Han, P. Hui, and A. Srinivasan, "Mobile data offloading in metropolitan area networks," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 14, pp. 28–30, November 2010.