How Much Can We Carry? A Capacity Analysis of Delay Tolerant Networking in Developing Countries

Abdullah Alhussainy
ETH Zurich
Gloriastrasse 35
Zurich
alhussab@tik.ee.ethz.ch

Karin Anna Hummel
ETH Zurich
Gloriastrasse 35
Zurich
hummel@tik.ee.ethz.ch

Panayotis Antoniadis
ETH Zurich
Gloriastrasse 35
Zurich
antoniadis@tik.ee.ethz.ch

ABSTRACT

In developing countries, where infrastructure data networks often provide only limited services, are too expensive, or are not available in wide parts, Delay Tolerant Networks (DTNs) provide a valuable complement for communications. Store-carry-forward or opportunistic networks have the potential to connect so far disconnected regions with regions where Internet access is granted. Yet, they depend on the existence of a critical mass of users acting as data carriers. The goal of this paper is to provide a way to estimate the capacity of such a network given available information on user mobility patterns at a countrywide scale. To do this, we analyze anonymized cellular mobility data of five months based on Call Detail Records (CDRs) provided by Orange for Ivory Coast. We derive inter-city and urban inbound and outbound flows of subscribers (carriers) and the corresponding estimated capacity of a hypothetical opportunistic network.

Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design—*Store and forward networks*

Keywords

Delay tolerant networks; mobility; capacity analysis

1. INTRODUCTION

Despite the constantly rising numbers of mobile device usage and the increased mobile Internet access, according to the ITU, still only 31 % of the population in developing countries can access the Internet [1]. Yet, a large number of mobile phones offer an interesting networking alternative as they can create a network by themselves, utilizing wireless ad-hoc networking technologies such as Bluetooth or Wi-Fi Direct. This network comes without cost on the infrastructure and follows the idea of resource pooling by the people [2]. In other words, the devices form opportunistic networks [3], i.e., networks built by mobile devices which store data locally, carry them while moving, and forward the data when coming into transmission range of another device. In this way, data of delay-tolerant applications may be transmitted by relaying through a few hops. When relaxing the requirement of fast delivery, data such as emails, podcasts, videos, and even tweets can be sent in a delay-tolerant fashion.

In this setting, the goal of this paper is to study the potential of opportunistic communications in order to extend the coverage of already existing infrastructure networks in terms of *capacity*. To make opportunistic communications more practical, we assume the existence of low-cost stationary storage or access points in populated areas, which we call *hubs*, for facilitating the data exchange between mobile devices and/or to provide access to the Internet (see also [4]). People moving between cities and villages may act as data carriers between these hubs. Our objective is to provide a first step to disclose the potential of opportunistic communications in developing countries. We present a simple model to derive the capacity of a hypothetical opportunistic network based on mobility flows derived from real traces. Our model includes additional influencing factors, such as the available storage space in mobile devices and the probability of a successful transmission for which we have assumed some intuitive values, which need to be updated through real life deployments in a next step.

To derive mobility flows, we analyze the D4D Challenge mobile phone data sets provided by Orange Ivory Coast, a major operator in Ivory Coast with a market share of about 25 % [5]. The anonymized data provide mobility information based on Call Detail Records (CDRs) of a subset of the customers of Orange over five months. Mobility information originating from the cellular network allows to infer physical movement on a large scale. Yet, using CDRs means to capture only mobile phones that are involved in phone calls or SMS messaging. The mobility flows we observe provide us with inidicative mobility patterns and lower bounds of flows. Yet, these data are a powerful source to give insights so far not possible for opportunistic networking aiming at connecting urban and rural regions.

We contribute with the following two sets of analysis:

- We provide a countrywide analysis of user mobility to derive spatial and temporal insights, that are, the popularity of locations, speed of users, and the distances covered by users (Section 3).
- We introduce a way to model the capacity of a hypothetical opportunistic network between major cities

LCDNet'13, September 30 - 29 2013, Miami, FL, USA

http://dx.doi.org/10.1145/2502880.2502889.

and urban and surrounding rural areas based on mobility flows derived from real cellular traces of Ivory Coast. Further, we consider the transmission success probability, storage capacity of mobile devices, and fraction of carriers capable and willing to participate. We discuss the implications of the model parameters on the feasibility of a video sharing application (Section 4).

Our preliminary results show that both inter-city and inbound and outbound urban mobility flows differ significantly among different cities. In case of Abidjan, one of the major cities of Ivory Coast, we show further variations for the different sectors of the surrounding rural area. Yet, given sufficient storage capacity on the mobile devices and transmission success probability, transferring a typical video to one of these sectors requires only a reasonable fraction of carriers participating in the opportunistic network.

2. RELATED WORK

The evaluation of opportunistic dissemination and routing protocols relies on mobility information about the carriers. Hence, various real world traces of different scales have been collected. Based on these traces, sequences of contacts are derived as well as inter-contact times and contact durations. The trace scales range from office scenarios, conferences, and campuses to city areas. Yet, to scale up nationwide is necessary to fully capture the potential for connecting so far disconnected regions in developing countries.

In addition to user traces, information about buses [6], planes [7], and vehicles [8] is used to either study the potential of the vehicles or the passengers acting as carriers. For example, in [7], airplane passengers, i.e., their mobile devices, act as carriers to transfer data between remote airports. The passenger flows are estimated based on the known aviation schedules. As intuitively expected, large data sizes benefit from the concept: a data size equal to the capacity of about three DVDs is recommended.

A source so far not often used by opportunistic networking research is the cellular network. Yet, more than four billion people on earth generate massive amounts of mobility related signaling in the cellular network. In particular the information stored in CDRs is a valuable source for studies, although the spatial accuracy is in principle limited to cell sizes. Yet, as the cellular network allows to estimate human mobility worldwide, it has been already used for studies on human mobility modeling [9] and flows within cities [10], etc. For example, Calabrese and Ratti [10] explicitly used mobile phone data to study the dynamics within cities related to urban planning policies. In the work of Eagle et al. [11], it is shown that rural and urban communities differ dramatically in terms of inferred behavioral characteristics. The investigated travels are based on CDRs of 1.4 million subscribers in a small country over four years. Wesolowski et al. [12] found that in Kibera, Kenya, about 50 % of the individuals are moving monthly either to other parts of Kibera or elsewhere. Such insights into user mobility are of particular interest for opportunistic networking which relies on mobility caused contacts. Additionally, studying the mobility and co-location of mobile devices allows to focus on exactly the part of the population that can potentially participate in an opportunistic network.

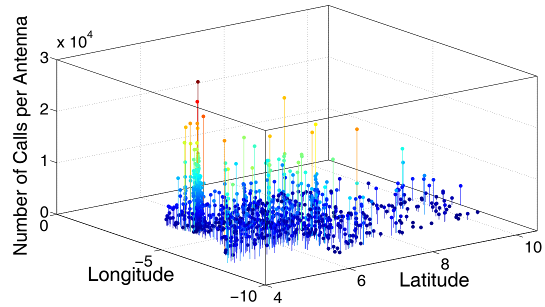


Figure 1: Number of calls per antenna (sample of two weeks).

Our study brings together opportunistic networking research with mobile phone data analysis. While the benefit of cellular mobility data has been already demonstrated in various domains, for DTNs, new questions arise concerning the capabilities of current or future mobile devices in use, the willingness of the people to participate, etc. We will address these valid considerations in the way we model the capacity of an opportunistic network.

3. MOBILITY ANALYSIS

We base our analysis of physical mobility in Ivory Coast on the mobility observed in a cellular network.

3.1 The Dataset

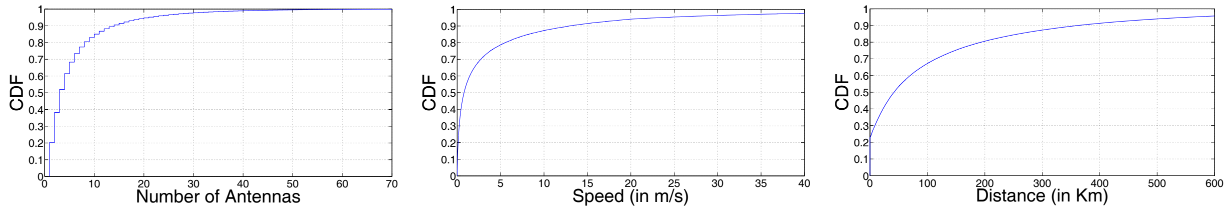
We use the D4D dataset [5] consisting of anonymized records originating from mobile phone calls and SMS between five million Orange customers in Ivory Coast, reported in the five-months period from December 1st, 2011 to April 28th, 2012. We selected the data sets with highest spatial resolution based on the calling activities of 50'000 randomly selected users, consisting of ten sample files, each covering about two weeks. The Caller and Antenna ID, and Date-Time are provided for every call and SMS sent. For privacy reasons, caller/user IDs are different among the ten sample files. Our results are based on the whole time frame of five months (ten two-weeks samples), unless stated otherwise.

3.2 From Cellular Data to User Mobility

An initial step in our analysis is to understand and visualize the mobility of users.

Calls per Area. The first spatial investigation we provide is on the number of users dwelling in areas. Therefore, we aggregate the number of calls per antenna and relate it to the geo-location of the antenna, illustrated for a single two-weeks sample in Figure 1. We observe that a higher number of calls (more frequently used antennas) indicates a more populated area, i.e., a city area, as expected. For example, the highest utilization was found for the very populated area of Abidjan. We will use this observation to define urban areas by a heuristic for our capacity calculation model in Section 4.

Speed and Distance. When a user is recorded to have made two consecutive calls through two different antennas, we consider it as a movement event. Without using geo-locations, the number of antennas connected to within a time frame is a first, simple measure of *speed*. Figure 2a shows the Cumulative Distribution Function (CDF) of the



(a) Antennas visited in two weeks.

(b) Speed.

(c) Distance covered in two weeks.

Figure 2: Mobility behavior of users.

number of antennas visited per user in two weeks. It can be inferred that 90 % of the users do not connect through more than 16 antennas in two weeks. The top active users in terms of the number of antennas could be in turn selected as potential data carriers.

A more detailed measure of speed can be provided by relating the geo-locations of the two antennas visited in sequence with the time advanced between the considered consecutive calls. Together with the geo-location of the antennas, we calculate the distance between antennas using the Haversine formula [13]. Given time difference and distance, we can further derive the traveling speed. Figure 2b depicts the CDF of the speed of the users observed when changing between antennas. We observe that 95 % of the users change antennas with speeds less than 25 m/s (90 km/h) which complies to the average speed limits in Ivory Coast, rural areas, that range between 80 and 120 km/h [14]. Note that the departure and arrival times are calculated based on the last time a call was made at the start position, and the first call made at the destination. The point of departure or arrival is the center of the coverage area of the corresponding antenna. The speeds calculated incorporate these inaccuracies.

The overall *distance* that a user covers in two weeks is calculated by accumulating the distances between antennas visited by the user consecutively (from the first to the last antenna visited). Figure 2c shows that 90 % of the users moved not more than about 363 km in two weeks. Note that this number includes commuting back and forth and movements between close antenna pairs which accumulate to the overall distance. We will now focus on isolating distant travels, which help us to estimate the capacity of an opportunistic network.

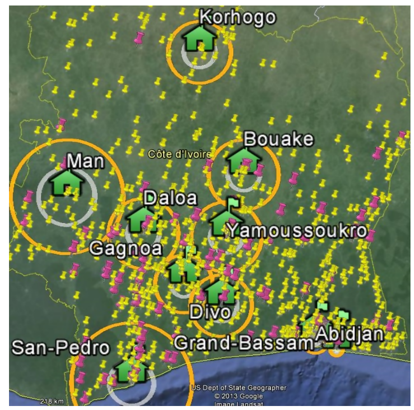


Figure 3: Selected cities and antennas: purple symbols represent the top 25 % most frequently visited antennas, yellow symbols the remaining antennas. The smaller grey circle defines the urban area, the larger orange circle defines the urban and surrounding rural area.

City name	Abbrev.	R (in m) Urban and surrounding area	r (in m) Urban area	Population [15, 16]
Abidjan	ABI	20'000	8'000	3'796'677
Bouake	BOU	55'000	27'500	659'233
Daloa	DAL	55'000	27'500	215'100
Yamoussoukro	YAM	55'000	27'500	200'659
Korhogo	KOR	50'000	25'000	142'039
Sanpedro	SAN	95'000	47'500	131'800
Man	MAN	90'000	45'000	116'657
Gagnoa	GAG	45'000	22'500	107'124
Divo	DIV	48'000	25'000	86'549
Grand-Bassam	G-B	10'000	5'000	83'576

Table 1: Ten major cities of Ivory Coast, their urban and surrounding area radii (R) and urban radii ($r \cong R/2$).

4. CAPACITY ESTIMATION

Our envisioned opportunistic network consists of human carriers transporting data on their mobile phones in a store-carry-forward fashion. The capacity of this network is determined by the mobility flow of the carriers as well as the storage capacity of the mobile phones. We introduce a formula for estimating the capacity of an opportunistic link between two locations, where a location is a geographic area.

We define the *capacity* c from location A to location B as the maximum amount of data that can be transported from A to B by the travelers in a time frame as follows (in bit/s):

$$c(A,B) = \frac{1}{T} \sum_{i=1}^{N_{\mathrm{A,B}}} d_i \cdot P_{i,\mathrm{success}}, \qquad (1)$$

where d_i is the storage capacity (in bits) of a mobile device i released for opportunistic communication, $N_{\mathrm{A,B}}$ is the number of travels from A to B in the time period T (in s), and $P_{i,\mathrm{success}}$ gives the probability that the data reaches the destination (within the time period T). The successful delivery of the data can be hampered due to inherent limitations of the available infrastructure such as a small number of hubs and insufficient coverage, or limited storage space per hub. Additionally, there are possible failures of hardware or software both at the hub system and user side to consider, such as energy outages. In our analysis, we will aggregate all these factors into the single value of $P_{i,\mathrm{success}}$, which for now we set intuitively leaving for future work the deployment of a real system in a developing country to make more realistic failure estimations. Finally, we extend the model by a

	BOU	DAL	YAM	KOR	SAN	MAN	GAG	DIV	G-B
ABI	36.7 (7.66)	20.30 (5.19)	136.95 (29.34)	5.65 (2.02)	22.4 (4.26)	11.25 (3.05)	11.25 (14.8)	66.85 (20.77)	418.4 (72.95)
BOU		6.80 (1.56)	223.95 (33.59)	38.1 (6.32)	1.95 (0.76)	1.15 (1.02)	4.4 (1.94)	2.9 (1.27)	0.35 (0.24)
DAL			48.55 (10.39)	0.75 (0.48)	7.55 (4.49)	17.15 (4.2)	6.60 (2.27)	1.15 (1.15)	0 (0)
YAM				7.8 (1.63)	3.25 (1.11)	3.1 (1.44)	17.7 (4.0)	4.35 (2.59)	0.7 (0.67)
KOR					0.2 (0.34)	0.45 (0.36)	0.25 (0.35)	0.15 (0.24)	0 (0)
SAN						2.65 (0.74)	9.55 (5.54)	1.95 (1.92)	0.45 (0.55)
MAN							6.35 (1.63)	1.85 (1.37)	0.05 (0.15)
GAG								69.2 (20.14)	0.60 (0.56)
DIV									0.7 (0.53)

Table 2: Inter-city flows per two weeks: average and standard deviation.

		ABI	BOU	DAL	YAM	KOR	SAN	MAN	GAG	DIV	G-B
Inbound	Average	36'095	895.50	1'087.04	1301.3	302.10	1'638.3	559.80	1'080.5	779.60	72.80
	Std. Deviation	6'713.9	217.49	233.9	195.54	107.48	291.88	134.75	209.41	402.64	20.05
Outbound	Average	36'315	909.10	1'106.9	1'281	308.50	1'646.3	559	1'071.9	687.90	78.40
	Std. Deviation	6'647.7	233.20	267.08	216.08	108.89	299.43	137.15	209.22	303.66	18.82

Table 3: Inbound and outbound flows per two weeks: average and standard deviation.

scaling factor in Section 4.4 and investigate the fraction of devices (owners) capable and willing to participate.

4.1 Model Assumptions

We model the mobility flow of humans concentrating on *inter-city* flows, here, flows between major cities of Ivory Coast, and on *inbound and outbound urban flows*, i.e., flows from urban areas to the surrounding villages (surrounding rural areas) and vice versa. To define the wider urban area consisting of the city and the surrounding rural area, we calculate the popularity of an antenna in terms of number of calls counted. Hereby, we assume that more calls are recorded in more populated areas, i.e., urban and surrounding areas. To be more precise, we select the n-quantile of antennas with highest number of calls recorded and cluster them along their geographic proximity, for $n = 25\,\%$. The clustering results in a circular wider urban area with radius R including the city and close suburban and rural areas. The radius R was slightly adjusted to avoid irrelevant intersections with areas of neighboring cities. The urban area is defined by another concentric circle with a radius of $r \cong R/2$ (again with some slight adjustments due to the concrete shapes of the urban and surrounding areas). Antennas that mainly belong to other cities than the ones we selected are not considered in our analysis. Figure 3 shows the ten urban and surrounding rural areas, details about the areas are given in Table 1.

Our main model assumption is that in each defined geographic area there exist storage and exchange points, which we term *hubs*. In case multiple hubs are deployed in one area, they are assumed to be interconnected. Mobile devices are expected to automatically start uploads and downloads according to the application when in proximity to a hub of the area. The number of hubs installed can increase the coverage, or, success of delivery of data carried to a specific area. Introducing such hubs to our model makes an opportunistic network more practical than assuming (solely) device to device communication. Thus, we consider direct phone to phone delivery as a possible complementary means of communications, yet, so far not considered in our capacity calculations.

Concerning mobility behavior, we assume that mobility is independent of hub placement, i.e., humans do not change their mobility behavior when hubs are deployed. Upon deployment of such an opportunistic network, we nevertheless expect that mobility will be affected in a way that will increase the capacity.

4.2 Population Flow between Areas

In the following, we detail inter-city flows as well as inbound and outbound flows. First, Table 2 details the inter-city flows. It can be observed that large, populated cities such as Abidjan or Yamoussoukro are central nodes, well connected to many other cities. As Table 3 shows for each city, inbound and outbound flows are – as expected – similar. Yet, they vary strongly for different cities. Abidjan shows the highest numbers of inbound and outbound flows compared to other cities. This is due to the large population of Abidjan of about 3.7 million inhabitants, highly populated surrounding suburbs such as Adobo (1'500'000 inhabitants) and Anayma (79'548 inhabitants), and reflects that Abidjan is an economically important area [16].

4.3 Capacity of the Opportunistic Network

The capacities to be expected from our hypothetical opportunistic network are calculated based on Eq.(1), according to which the capacity is proportional to the number of flows and the storage capacity of the mobile phones, and can be reduced by the delivery success probability. The latter parameters are subject to future investigations and can be easily used to model heterogeneous mobile devices and various failures, respectively. In the following, we give capacity calculations for sample settings, i.e., an available storage capacity of $d = 100$ MB on each device and a pessimistic delivery success probability $P_{\text{success}} = 0.2$. For inbound and outbound capacity, we slice the whole surrounding area into *sectors* of about the size of a village, i.e., about $50\,\text{km}^2$.

For inter-city capacities, Figure 4 shows that the highest capacities are achieved in a range of hundreds of kbit/s. Larger cities such as Abidjan and Yamoussoukro show higher capacity links to cities in proximity. For example, the highest capacity of about 55 kbit/s is observed for Abidjan and Grand-Bassam which are about 42 km apart.

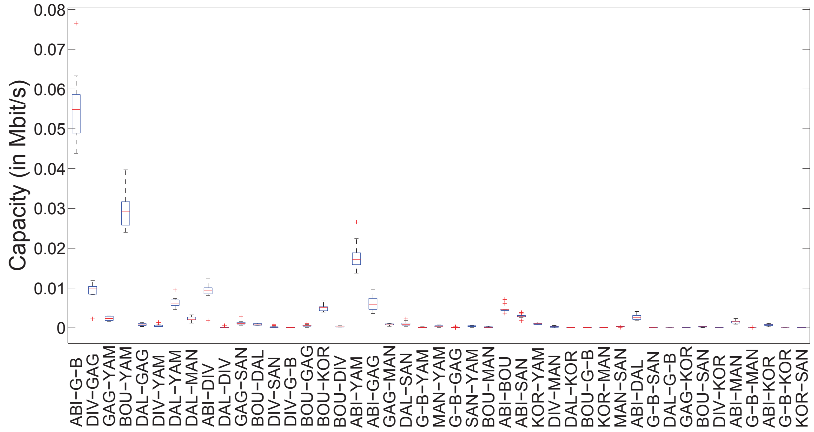


Figure 4: Capacity of city pairs: boxplots give the median (red line in the box), other quartiles (box), and extreme values of ten two-weeks flow samples ($d = 100$ MB, $P_{\text{success}} = 0.2$).

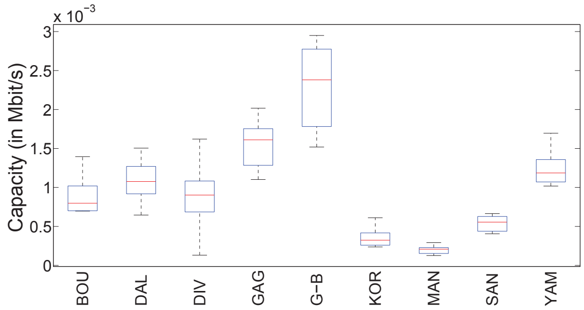


Figure 5: Average outbound capacity per $50\,\text{km}^2$ sector of nine major cities of Ivory Coast: boxplots as in Figure 4 ($d = 100$ MB, $P_{\text{success}} = 0.2$).

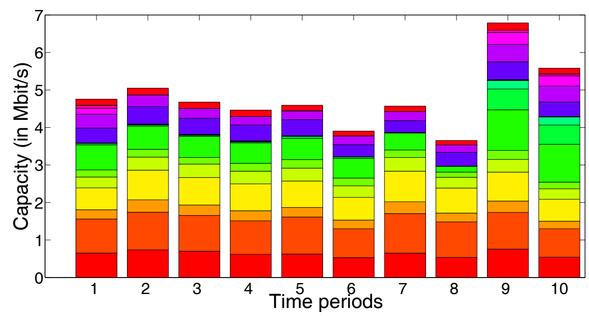


Figure 6: Outbound capacity of Abidjan for each of the two-week periods: Elements of each bar represent the outbound capacity of one sector ($d = 100$ MB, $P_{\text{success}} = 0.2$; sectors that have no antennas are omitted).

Figure 5 shows the average outbound capacity per sector of nine major cities in Ivory Coast (not shown here, the inbound capacity, which is similar). The highest capacity is observed for Grand-Bassam, i.e., about 2.4 kbit/s (given the pessimistic setting of $P_{\text{success}} = 0.2$). The urban outbound capacity of the most active city, Abidjan, is detailed in Figure 6, broken down to the ten two-week time periods of the data set and stacked for the sectors. The two-weeks samples vary in terms of overall capacity, but also sector-wise. In week nine, the maximum outbound capacity reached approximately 6.6 Mbit/s (about 0.44 Mbit/s per sector on average).

The calculated sample capacity values consider, on the one hand, mobility flows derived from real traces, and, on the other hand, sample parameter values for d and P_{success} which are subject to adoption to real life observations. The capacity will increase with increasing d and P_{success} (see Eq.(1)).

4.4 Use Case: Video Sharing

The capacity values derived so far summarize the outcome for the provided dataset, i.e., a selected subset of customers of Orange in Ivory Coast (50'000 customers out of overall five million customers [5]). If we assume that the sample is a representative one, we can estimate the potential capacity achieved by a larger number of people.

The basic problem formulation in Eq.(1) is now extended to include a scaling factor s in order to scale up to a larger commuter set ($s \geq 1$), and a factor f ($0 \leq f \leq 1$) describing the fraction of the users who are willing and capable to participate as carriers. The capacity is now calculated as:

$$c_{\text{ext}}(A, B) = s \cdot f \cdot c(A, B). \tag{2}$$

When considering the five million customers of Orange in Ivory Coast [5], the scaling factor is $s = 100$. To give an idea whether the capacity per sector is sufficient to successfully support an Internet application, we investigate the fraction f of mobile carriers (each using 100 MB of storage to carry the data) required to deliver a Youtube video of, e.g., 500 MB (e.g., a documentary on sustainable energy, how to avoid HIV infection, farming, etc.) from Abidjan to a sector of the surrounding rural area. Hereby, we are only interested in when the video is available at an area, not when it is fetched by a single user. The application envisioned is the delivery of popular videos of interest for many people. Again, we set $d = 100$ MB, and $P_{\text{success}} = 0.2$. We present below our calculation for Abidjan with a surrounding rural area of about 1'055 km^2, the sector area is again about $50\,\text{km}^2$ (resulting in 20 sectors). We first observe, that the 153 antennas located in the surrounding rural area are not uniformly distributed among the sectors.

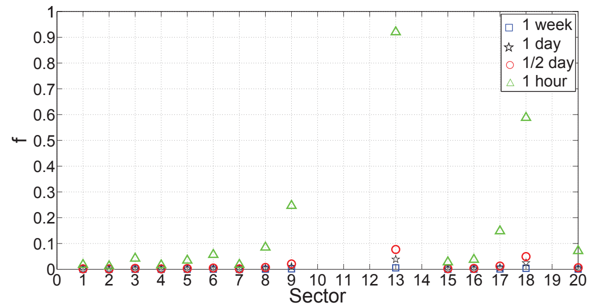


Figure 7: Use case video sharing: fraction f of users required to transfer a 500 MB video from Abidjan to each surrounding sector ($d = 100$ MB, $P_{\text{success}} = 0.2$, $s = 100$).

Then, we introduce a maximum tolerated delay t for the transmission, or, in other words, the following condition is stated for the required capacity: $c_{\text{ext}} > \frac{500 \cdot 8 \cdot 10^6}{t}$ (bit/s). Figure 7 shows the fraction f required for delivering the video with different tolerated delays detailed for each sector. No flows are observed for five sectors, among the others, most sectors require a fraction f in the range of below 1 % to transmit the video in one hour. Only a few sectors require a higher fraction, e.g., a video transferred from a city Internet hub at Abidjan can be transferred to Sector 13 within one hour, if about 91 % of the users traveling will accept to carry the data (24 individuals). If we assume a more optimistic delivery success rate and set $P_{\text{success}} = 0.8$, the fraction of users decreases to about 23 % (six individuals). Note that f decreases linearly with the increase of the delay.

5. DISCUSSION

To capture the mobility of people in Ivory Coast, we used Orange's D4D cellular dataset. Then, based on a simple formula that introduced major factors of an opportunistic network, we estimated the capacity of such a hypothetical network, and analyzed under which conditions a video sharing application can be supported in the Abidjan area.

Clearly, our capacity calculation is indicative since it depends on a dataset from which only a subset of the actual mobility of people can be inferred. Further, the feasibility of our opportunistic scenario depends on the proliferation of advanced mobile devices like smartphones with sufficient storage capacity and the willingness of the population to cooperate, two of our model's parameters. Finally, we assumed the existence of a certain number of low-cost hubs that can offer Internet connectivity or just facilitate data exchange by being placed in appropriate locations in cities and large villages. We believe that this is an important prerequisite for the success of an opportunistic network. Yet, we leave the analysis of the effect of such hubs on current mobility patterns and the optimal placement of hubs by studying, e.g., inner-city movement, to future studies. We have further introduced several parameters to the capacity calculation, which we plan to investigate further as well as to extend our analysis to more datasets and more candidate applications.

6. ACKNOWLEDGMENTS

Studies and researches were performed using mobile communication data made available by France Telecom and Orange Ivory Coast within the D4D Challenge[1]. This work has been supported by the SNSF and the North-South Center, ETH Zurich, by the EC under the FP7 contract PIEF-GA-2010-276336 MOVE-R and by EINS, the Network of Excellence in Internet Science, FP7 grant 288021.

7. REFERENCES

[1] The World in 2013, ICT Facts and Figures. Technical report, Geneva, Switzerland: ITU, 2013.

[2] A. Sathiaseelan and J. Crowcroft. LCD-Net: Lowest Cost Denominator Networking. *SIGCOMM Comput. Commun. Rev.*, 43(2):52–57, Apr 2013.

[3] M. Conti and M. Kumar. Opportunities in Opportunistic Computing. *Computer*, 43(1):42–50, 2010.

[4] Z. Wenrui, Y. Chen, M. Ammar, M. Corner, B. Levine, and E. Zegura. Capacity Enhancement using Throwboxes in DTNs. In *IEEE International Conference on Mobile Adhoc and Sensor Systems (MASS)*, pages 31–40, 2006.

[5] V. Blondel, M. Esch, C. Chan, F. Clerot, P. Deville, E. Huens, F. Morlot, Z. Smoreda, and C. Ziemlicki. Data for Development: the D4D Challenge on Mobile Phone Data. *arXiv preprint arXiv:1210.0137*, 2012.

[6] X. Zhang, J. Kurose, B. Levine, D. Towsley, and H. Zhang. Study of a Bus-based Disruption-tolerant Network: Mobility Modeling and Impact on Routing. In *13th ACM Int. Conf. on Mobile Computing and Networking*, pages 195–206, 2007.

[7] A. Keränen and J. Ott. DTN Over Aerial Carriers. In *4th ACM Workshop on Challenged Networks*, pages 67–76, 2009.

[8] R. Gorcitz, Y. Jarma, P. Spathis, M. Dias de Amorim, R. Wakikawa, J. Whitbeck, V. Conan, and S. Fdida. Vehicular Carriers for Big Data Transfers (Poster). In *IEEE Vehicular Networking Conf.*, pages 109–114, 2012.

[9] M.C. Gonzalez, C.A. Hidalgo, and A.L. Barabasi. Understanding Individual Human Mobility Patterns. *Nature*, 453(7196):779–782, 2008.

[10] F. Calabrese and C. Ratti. Real Time Rome. *Networks and Communication Studies*, 20(3-4):247–258, 2006.

[11] N. Eagle, Y.A. de Montjoye, and L.M.A. Bettencourt. Community Computing: Comparisons between Rural and Urban Societies Using Mobile Phone Data. In *Int. Conf. on Computational Science and Engineering*, pages 144–150, 2009.

[12] A. Wesolowski and N. Eagle. Inferring Human Dynamics in Slums Using Mobile Phone Data. Technical report, Santa Fe Institute, 2009.

[13] C.C. Robusto. The Cosine-Haversine Formula. *The American Mathematical Monthly*, 64(1):38–40, 1957.

[14] T. Assum. Road Safety in Africa: Appraisal of Road Safety Initiatives in Five African Countries. *Sub-Saharan Africa Transport Policy Program. Working Paper*, 33, 1998.

[15] International Census. `http://www.census.gov/population/international/`.

[16] K. Appessika. The Case of Abidjan. *Ivory Coast, Case Study for the Global Report*, 2003.

[1] D4D Challenge: `http://www.d4d.orange.com/`