Human Mobility Support for Personalised Data Offloading

Emanuel Lima, Ana Aguiar, *Member, IEEE*, Paulo Carvalho and Aline Carneiro Viana

***Abstract*—WiFi Access Points (APs) can be used to offload data or computation tasks while users are commuting. However, due to APs' limited coverage, offloading performance is heavily impacted by the users' mobility. This work proposes to leverage human mobility to inform offloading tasks, taking a data based approach leveraging granular mobility datasets from two cities: Porto and Beijing. We define Offloading Regions (ORs) as areas where a commuter's mobility would enable offloading, and propose an unsupervised learning methodology to extract ORs from mobility traces. Then, we characterise and analyse ORs according to offloading opportunity metrics such as type, availability, total time to offload, and offloading delay. Results show that in 50% of the trips, users spend more than 48% of the travel time inside ORs extracted according to the proposed methodology.**

The ability to predict the next ORs would benefit offloading orchestration. Offloading mobility predictability, although crucial, proves to be challenging, expressed by the poor predictive performance of well-known models ($\approx$ 37% acc. for the best predictor). We show that mobility regularity properties improve predictive performance up to $\approx$ 35%.

Finally, we look into the impact of further OR extraction and prediction parameters. We show that the exploration phase length does not impact the discovery of low relevance ORs, and that both filtering low relevance OR and predicting multiple ORs increase predictability. By characterising the trade-off between mobility predictability and offloading opportunities in transit, we highlighting the need for offloading systems to adopt hybrid strategies, i.e., mixing opportunistic and predictive strategies. The conclusions and findings on offloading mobility properties are likely to generalise for varied urban scenarios given the high degree of similarity between the results obtained for the two different and independently collected mobility datasets.

***Index Terms*—Data Offloading, Human Mobility, Mobility Predictability, Offloading Mobility Properties, Offloading Systems**

I. Introduction

MOBILE data traffic has been growing tremendously with the increasing of the number of applications leveraging ubiquitous Internet access. According to Cisco forecasts [1], mobile data traffic is growing at a compound annual growth rate of 46% since 2017 and will continue to increase. Mobile network operators are struggling to keep up with this traffic demand, and part of the solution is to offload communications to WiFi networks [2]. The massive deployment of WiFi networks in homes and businesses is increasing the Internet coverage in densely populated urban areas, making these networks attractive to offload mobile data or access computation resources in the cloud or at the edge [3].

It is well known that most people spend their time in specific locations, known as personal Points of Interest (PoI), e.g., home, work or favorite restaurant, which usually have WiFi networks that are used to offload data. Therefore, it can be argued that cellular networks overload will be mainly caused by users while travelling since during these periods the cellular network is commonly the default option for Internet access. In such scenario, deferring data transmissions while the user is travelling, offloading only at the trip's destination, was shown to be a feasible solution [4]. However, most applications cannot sustain large delays without impacting the user's quality of experience (QoE). For this reason, it is crucial to develop mobile offloading systems that proactively take decisions on WiFi places and connectivity windows to offload data or tasks, while the users are travelling. In order to do that, mobile offloading systems must rely on detailed information regarding the users' mobility to assist mobile devices deciding when and where to offload.

This work is the first to explore mobility data from the mobile offloading perspective, focusing on periods where the users are in transit. We argue that mobile offloading systems need to rely on offloading mobility profiles, which provide detailed information about the users' mobility, to assist mobile devices in deciding when and where to offload to WiFi networks. We use mobility data to show that such a strategy is necessary and feasible to accommodate the differentiated user connectivity profiles during commute. This sort of strategies will run on the user's device and, after learning their habits of movement, will adapt and extract the best regions to perform offload. This brings a perceptive feature to the offloading strategy and the proposed learning methods are general to any type of dataset.

Our paper explores how human mobility impacts mobile offloading systems during commuting trips bringing evidence that individual profiles should be considered. We seek answers to questions like:

1) How can individual commuter offloading mobility profiles be established?

This work is a result of the project FLOYD (POCI-01-0247-FEDER-045912), funded by the European Regional Development Fund (FEDER), through the Operational Competitiveness and Internationalization Programme (COMPETE 2020) and by National Funds (OE), through Fundação para a Ciência e Tecnologia, I.P.; and UIDB/50008/2020 funded by the applicable financial framework (FCT/MCTES) (PIDDAC). It also has been supported by FCT – Fundação para a Ciência e Tecnologia within the RD Units Project Scope: UIDB/00319/2020 and is related to the Hi!Paris project entitled "AI for More Accessible Cities"

The authors Emanuel Lima and Ana Aguiar are with Instituto de Telecomunicações, Department of Electrical and Computer Engineering from the Faculty of Engineering of the University of Porto, Portugal (e-mail: {emanuellima;anaa}@fe.up.pt). The author Paulo Carvalho is with Centro Algoritmi, Departamento de Informática from Universidade do Minho, Braga, Portugal (e-mail: pmc@di.uminho.pt) and Aline Viana is with Inria, Palaiseau, France (e-mail: aline.viana@inria.fr).

2) Does human mobility in urban areas allow for offloading to limited coverage technologies?
3) What are the limits of the predictability of per user offloading connectivity and which factors impact that predictability?
4) Which trade-offs should be considered in the system design?

To answer this questions we develop and implement the methodology pipeline presented in Figure 1. This methodology encompasses a method to establish individual offloading mobility profiles that would be part of the offloading strategies we propose — *system components*; and the steps that provide quantitative answers on the feasibility of such a system — *system evaluation*. We use a data based approach relying on two granular datasets from Porto and Beijing to support the generality of the results. Therefore, we bring answer to the first question in our contributions (i) and (ii). We answer the second and the third question in the contributions (iii) and (iv) respectively, and finally, we answer the last question in the contribution (v). In short, we define our contributions as follows:

(i) We propose an unsupervised learning methodology to identify individual offloading regions (ORs), reflecting user habits and routines, considering different time windows to offload. Details are given in Section IV.

(ii) In Section V we categorize ORs in terms of their relevance and spatial characteristics. A comprehensive evaluation shows that although collected in different cities and time periods, the laws driving users' mobility reflect similar relevance and spacial characteristics of the ORs for both datasets.

(iii) We evaluate the offloading opportunities offered to users at the OR while travelling in terms of type, availability, time to offload and offloading delay. Our results show the small offloading delays along with the high-temporal coverage of ORs confirm opportunities to offload to limited coverage cells while the users are in transit. Details are given in Section VI.

(iv) We assess the mobility predictability in an offloading scenario using theoretical and algorithmic evaluation of several mobility predictors. The results show that mobility predictability for offloading purposes is far more challenging than mobility between PoIs. Here, machine learning (ML) predictors outperform common Markov Chain (MC) predictors used in the literature by at least 15%, revealing the importance of context information in an offloading scenario. Details are given in Section VII.

(v) Finally, we propose and discuss further flavours that could impact OR extraction and predictability, namely the *exploration phase length*[1], the *mobility regularity*, and *offloading locations*, as well as their impact on the design of offloading systems. Specifically, we identify that considering longer mobility learning periods is unlike to improve the capacity to predict offloading mobility. However, mobility regularity can be leveraged to improve predictability by $\approx 27\%$ at the expense of fewer offloading opportunities. Thus, offloading systems should rely on a combination of opportunistic and deterministic strategies. Attending to the characteristics of offloading locations observed in both datasets, we show that in the majority of the cases, APs already deployed in urban environments can provide full coverage to users while offloading, decreasing the need for handovers. Details are given in Section-VIII.

[1]Exploration phase is the time to visit each OR for the first time.

The remainder of this paper is organized as follows. The related work on human mobility and mobile offloading is discussed in Section II. The datasets used in this work are described in Section III while the methodology to extract ORs from trajectory traces is introduced in Section IV. The characteristics of ORs are presented in Section V and the analysis of the offloading opportunities offered to users in transit is detailed in Section VI. The mobility predictability in an offloading scenario is studied in Section VII. Finally, the different perspectives of our work are debated in Section VIII, and the main conclusions presented in Section IX.

II. Related Work

Mobile offloading can be infrastructure-based or infrastructure-less depending on the first-hop type during the offloading process. In the first, mobile devices offload directly to APs. Instead, in the infrastructure-less approach, data is sent to other mobile devices, and the offloading results from the combined mobility of a sequence of non-controllable entities [5], [6]. In this work, *we focus on infrastructure-based delayed offloading as it has been proven to reduce congestion in cellular networks considerably at the expense of small transmission delays* [4], [7]. Here, several models have been proposed to improve the offloading decision [8], [9], which dictate "when" to offload considering variables such as traffic constraints, offloading delay, efficiency, volume, etc. However, before deciding "when" to offload, we first need to investigate how good are the moments and places where devices are in range of WiFi infrastructure. This depends on the offloading opportunities leveraged from human mobility, and thus mobility behavior and predictability of individuals.

Human mobility has been mostly focused on two main domains: the identification and characterization of locations as users' personal PoI and the study of mobility patterns between these locations. Several other works have been proposed to infer PoI from GPS mobility traces [10], [11]. Usually, PoI are associated with a staying time which confers its degree of importance, commonly in the order of several minutes. This time is suitable for detecting important locations but not for inferring offloading occasions, where short contacts with WiFi APs can be leveraged for offloading.

Human mobility between PoI has been shown to be very predictable. In [12], a theoretical framework is proposed and shows that the upper limit for the predictive performance on predicting the next PoI visited by a user is surprisingly high (93%), staying constant throughout heterogeneous sets of users (e.g., for different gender, age, geographical attachment). When spatial reachability constraints are considered, a tighter upper bound of 81-85% can be obtained [13]. A large diversity of mobility predictors using Recurrent Neural Networks [14], Bayes models [15], Random Forest [16], and

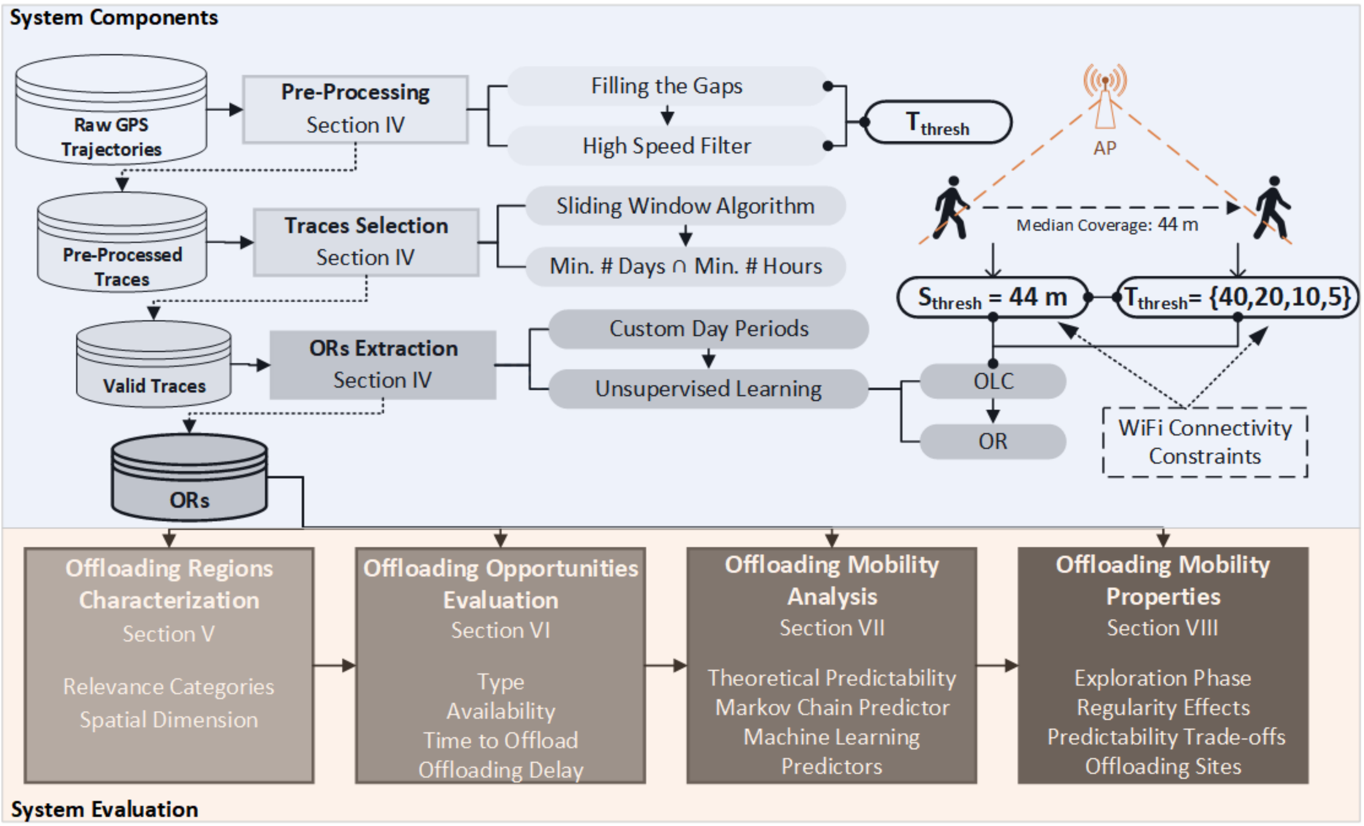


Fig. 1: Offloading mobility analysis methodology.

Markov Chains [17], [18] has been proposed, approaching the theoretical predictability. However, most of these studies focus on mobility between personal PoI and not on scenarios of offloading opportunities while users are travelling between these locations.

In [19], rather that focusing on PoI, we first introduce a methodology to identify and extract offloading zones from individual GPS trajectories, when small offloading time windows are considered. Despite the methodology proposed, the mobility patterns between offloading zones are not explored and the study was conducted on a single dataset. In the present work, we use the same methodology to extract offloading zones from two mobility datasets collected in different countries at different times and analyse the users' mobility in an offloading scenario. To the best of our knowledge, this is the first work to study and characterize human mobility for the purpose of offloading when the users are in transit.

III. Mobility Datasets

Several data sources have been used to study human mobility but high sampled GPS trajectories remain the most reliable way to track outdoor movements. The main dataset used in this work was collected in the city of Porto, Portugal (Figure 2a) from 2016 to 2018 through several crowdsensing campaigns, and involved a total of 408 users. During these campaigns, GPS data, i.e, user's speed, estimated user location (latitude, longitude), and the accuracy of this estimation, were collected at a high temporal resolution (see Table I). Moreover, automatic start/stop strategies were being used by the application to detect the beginning and the end of a trip. At the end of each trip, a pop-up with a survey was automatically presented to the users asking for movement confirmation. This data collection process is detailed in [20] and provides two unique features to the dataset: (i) high temporal resolution of positioning data of 1Hz; and (ii) annotated mobility traces concerning the beginning and the end of a trip.

In order to ensure that the results presented in this work are not biased by the usage of a single dataset, a second dataset, namely Geolife [21], is used. This dataset provides time-stamped GPS locations of 182 individuals collected from 2007 to 2012 and its choice was due to two main factors: (i) it is widely used in human mobility studies as it is one of the first publicly released datasets containing mobility; and (ii) the trajectories are logged in a dense representation – 91.5% of the trajectories are logged every 1-5 seconds or every 5-10 meters per point. As most data from Geolife was collected in Beijing, only mobility traces collected in the metropolitan area of this city (area in Figure 2b) are considered in this work, comprising data from a total of 133 users, Table I.

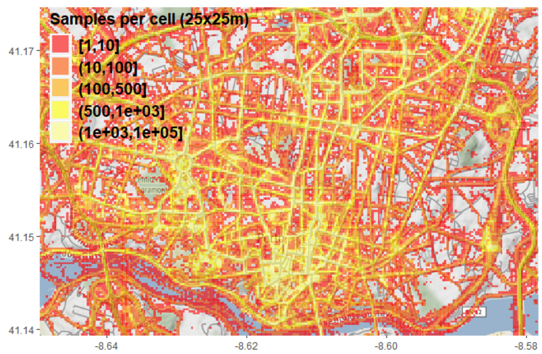
Both Porto and Beijing datasets include a similar amount of location samples ($\approx 10M$), however, as shown in Figure 2, due to the different cites' area, the dataset from Porto presents a higher density of location samples per cell (25x25m).

Although there are much more users in Porto's dataset when inspecting the number of weekdays with data per user, Figure 3 shows that users from Beijing present almost six times more

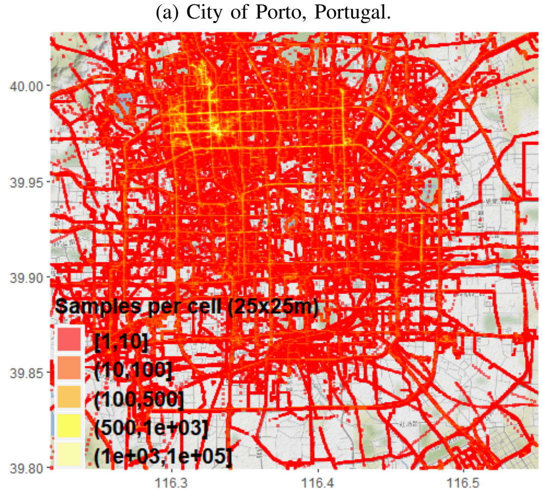
TABLE I: Datasets main characteristics.

	Porto	Beijing
Period	2016-2018	2007-2012
Users	408	133
Area km^2	24	723
Locations Samples	9.7M	10M
Temporal resolution	1s-2s	1s-5s

days with data. While the median number of weekdays for the Porto's users is 5 weekdays, in Beijing, is 28 weekdays. However, when focusing on weekends, both datasets present a median number of 2.5 days per user. The large difference between the number of weekdays and weekends in Beijing dataset is mostly due to the fact that most of users lived outside Beijing metropolitan area (Figure 2b), therefore, not traveling to the city center on weekends.



(a) City of Porto, Portugal.



(b) City of Beijing, China.

Fig. 2: Heatmap of collected location samples per 25x25m cells.

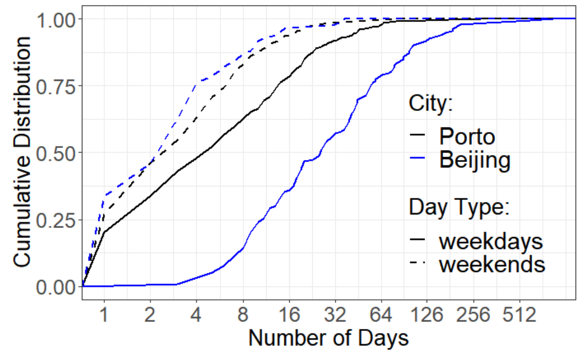


Fig. 3: Number of days with data per user.

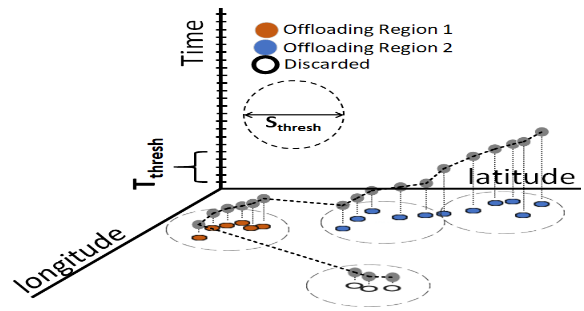


Fig. 4: ORs extraction from a user's trajectory [19].

IV. Inferring Offloading Regions

Offloading Regions (ORs) are defined as geographical areas where a user exhibits a mobility suitable for offloading. The notion and the methodology for extracting ORs from trajectory traces were initial introduced in our previous work [19]. In this section, to keep this article self-contained, we summarize the key concepts and the methodology for extracting ORs from trajectory traces and start answering the first question presented in the introduction.

A. Mobility Constraints for Offloading

Due to the limited coverage of WiFi APs, users' mobility dictates the time window available for offloading. Locations where a user is stopped or moving at low speeds are preferred for offloading, as the time window is maximized. Thus, ORs are identified by applying space and time constraints to the users' mobility. We define the spatial threshold S_{thresh} as the maximum distance between two points that can be considered to be in the same offloading opportunity. This threshold should reflect the coverage of the technologies used. And we define T_{thresh} as the minimum time that a user must spend in a geographical area defined by the S_{thresh} to be able to take advantage of offloading. Therefore, in order to define an OR, we introduce the concept of offloading location candidate (OLC). An OLC is a geographical area defined by S_{thresh}, where a user stays during T_{thresh}. Then, an OR is defined as the aggregation of contiguous OLCs (see Figure 4), reflecting aggregated areas where the user has a mobility suitable for offloading, e.g. walking path from home to bus station. When no aggregation can be performed, an OR corresponds to an OLC. Section IV-D brings more details on the ORs' extraction methodology.

Considering an offloading scenario based on WiFi networks, the coverage of WiFi APs within an urban environment is used as space threshold. Ground truth from more than 40k APs shows that the median coverage of a AP in an urban scenario is 44m [19]. Thus, we use this value for S_{thresh}. The time threshold should reflect that different applications may require different minimum time windows for offloading, and that some offloading orchestration time is necessary. In order to discover opportunities that could use short time windows to offload as well as less mobility friendly configurations, this study considers ($T_{thresh} = \{40, 20, 10, 5\}$), Figure 4.

B. Trajectories Pre-Processing

The first stage of ORs' extraction is pre-processing, as shown in Figure 1. The users' trajectories may include gaps from periods where no GPS data was collected. This may be due to lack of GPS signal in indoor environments or mobile device/application being turned off. To avoid the underestimation of ORs and to capture locations where users were stopped or inside buildings, we identify the gaps G corresponding to these cases using a $max(G_{dist}) = 100m$ (cf. Definition 1). Then, for the whole duration of G (in sec), we use linear interpolation to fill these gaps by adding pseudo locations.

Definition 1: Let a user's trajectory for the user u be a set of locations $L_i^u = (latitude, longitude)$ collected at the timestamp i represented as $Traj^u = (L_i^u, L_{i+1}^u, ..., L_{i+n}^u)$. We define G as the spatio-temporal break satisfying the condition $(distance(L_i^u, L_{i+1}^u) < G_{dist}) \wedge (time(L_i^u, L_{i+1}^u) > T)$, for $T = \frac{G_{dist}}{max\_speed}$ where max_speed is given by Eq. 1.

As ORs are defined by T_{thresh} and S_{thresh}, the users' speed is limited by:

$$max\_speed = \frac{S_{thresh}}{T_{thresh}} \tag{1}$$

Then, we use a high-speed filter for eliminating from the traces samples not satisfying Eq. 1.

C. Mobility Traces Selection

In the second stage of Figure 1, a filter is also applied to the mobility traces in order to select users that provide enough historical data for studying their mobility patterns. As there is a reduced number of weekends in both datasets (see Figure 3), only mobility traces collected from weekdays are considered.

Due to the extensive data collection period in both datasets, some users exhibit mobility traces that were collected several months or even years apart. Traces with such a large time offset may lead to wrong assumptions regarding the user's mobility behaviors. Therefore, a time sliding window of 4 months is defined and applied to each user in order to select consecutive months where the user presents the highest number of days with data in the dataset. Users are then selected based on two filters: (i) minimum number of days, and (ii) minimum number of hours with, at least, one GPS sample. Figure 5 represents the number of users after intersecting these two filters and shows that both datasets present a similar number of users with a large number of hours per day with GPS (top left corner of the heatmaps). However, as the minimum number of days threshold increases, Beijing dataset provides more users with a high number of GPS hours per day (top right corner of the heatmaps). This reveals a higher degree of engagement of users in Geolife towards the data collection process when compared to users in Porto.

To avoid favoring a single dataset, we only consider traces from users having at least 5 weekdays with 8h of GPS data, resulting in mobility traces from a total of 55 and 57 users in Porto and Beijing datasets, respectively.

TABLE II: Activity periods in a day.

Definition	Day Period	Usual Activities
DP1	01:00h to 06:59h	Sleeping
DP2	07:00h to 09:59h	Commuting
DP3	10:00h to 11:59h	Working
DP4	12:00h to 13:59h	Lunch
DP5	14:00h to 16:59h	Working
DP6	17:00h to 19:59h	Commuting
DP7	20:00h to 00:59h	Social

D. Offloading Regions Extraction

Finally, in order to extract ORs from the users' trajectory traces the DBSCAN [22] clustering algorithm is used for the identification and aggregation of OLCs. The use of a density-based clustering algorithm to extract ORs allows to obtain clusters with arbitrary shapes, focusing only on locations/areas with high concentration of points. There are two parameters that need to be set in DBSCAN before the clustering process: *MinPts* and ϵ. A data point is a core point if it has at least *MinPts* in its neighborhood ϵ. This notion of core point is adopted to define an OLC. As in our dataset the location samples are produced at a fixed rate of 1 Hz, *MinPts* is defined as T_{thresh} and ϵ as S_{thresh}, once more reflecting the WiFi connectivity into spatial constraints. Then, ORs are the clusters comprising OLCs that are density connected. Clusters not meeting the OLC criterion are discarded, while adjacent OLCs are aggregated in the same cluster, forming ORs.

The analysis is focused on weekdays considering periods that usually represent daily routines in people lives, as shown in Table II. Therefore, for each user, the clustering process described above is applied to the trajectories in each period. In the downstream analysis, we assume that inside an OR, there is contiguous and continuous connectivity adequate for offloading data or computation.

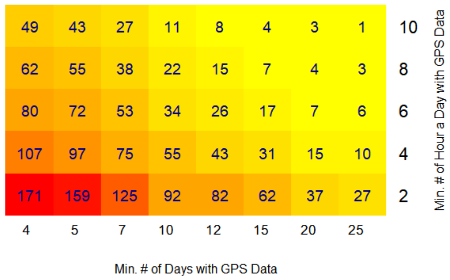
V. Characterizing Offloading Regions

ORs represent locations where offloading tasks can be performed. In this section, these locations are characterized in detail as the first stage to characterize offloading mobility and continue to answer the first question presented in the introduction.

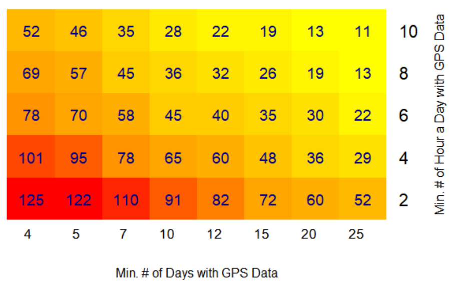
A. Categories of Relevance

In order to better characterize ORs, we classify them into categories of relevance. *Relevance* of OR x is defined based on the frequency of user u visiting OR_x, i.e., $R_{OR_x}^u = \frac{D_{OR_x}^u}{D_{DP}^u}$, where $D_{OR_x}^u$ is the total number of days that OR_x was visited by the user u and D_{DP}^u is the total number of days in the dataset for the user u within the day period DP (Table II).

The number of relevance categories that best suits all users is estimated using unsupervised learning. In this way, k-means algorithm is applied to each user relevance values for a different number of clusters (k), where k represents the relevance categories in use. The optimal k is then determined using the elbow method along with the total WSS (Within Sum of Squares) [23]. For the vast majority of users, k=3 gives the best number of clusters, which determines three categories of relevance: *low, medium*, and *high*.

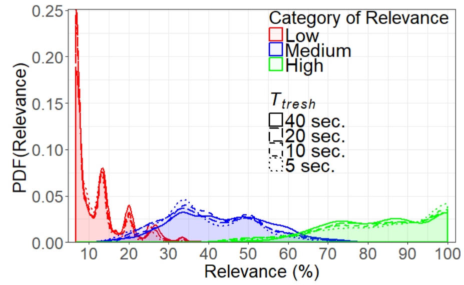


(a) City of Porto, Portugal.

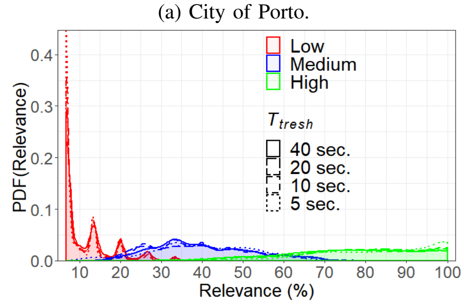


(b) City of Beijing, China.

Fig. 5: Number of users as the intersection of number of days and hours filters.

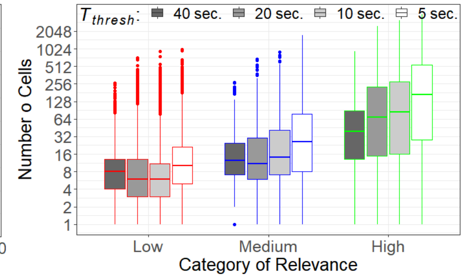


(a) City of Porto.

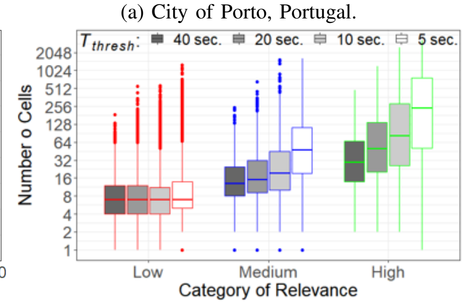


(b) City of Beijing.

Fig. 6: ORs relevance values per category.



(a) City of Porto, Portugal.



(b) City of Beijing, China.

Fig. 7: ORs' number of cells (11x11m).

Mapping ORs to these categories implies the definition of relevance values to be used as bounds for each category. As these bounds depend on the users' mobility patterns, they cannot be predefined as each user has its own notion of relevance. To solve this problem user-base thresholds are needed, and therefore, the ORs relevance values of each user are clustered using k-means with *k=3*, and then, each cluster is classified as *low*, *medium* or *high* category according to the distribution of relevance values.

Figure 6 shows the probability density function of the ORs' relevance for all users, obtained using the kernel density estimation for both cities. The clean separation between the categories' distributions validates the choice of the categories of relevance (*k=3*). Besides, the small overlapping area between distributions reveals that most users have similar bounds for each relevance category. However, the fact that there is an overlap shows the importance of a user-base threshold approach for the ORs' relevance classification, as an OR with 25% relevance may be low-relevant to a user, while other OR with the same relevance value may be medium-relevant to other user. Moreover, *it can be seen that users from Porto and Beijing exhibit similar relevance distribution concerning the categories of relevance. This reveals that ORs' visiting behaviors appear not to depend on the country or city but instead on the mobility pattern features in the offloading scenario.* ORs within the high-relevance category are visited by users almost daily with the distribution peaking at 100%. The high-relevance category includes locations such as home, work place, and commute paths. The ORs within the medium-relevance category may be occasionally visited by users (favorite restaurant, gym, etc.) with the distribution peaking at $\approx$ 33%. Finally, ORs within the low-relevance category are visited sporadically (distribution peaking approximately at 7%, 14%, and 20%).

B. *Spatial Characteristics*

As ORs can be used as offloading sites, their spatial characteristics represent the areas where offloading can be performed. Figure 7 shows the distribution of the number of cells (11x11m grid squares) in the ORs for each category of relevance. As illustrated, the size of ORs increases with the relevance category for both datasets. The median number of cells when considering a $T_{thresh} = \{40, 20, 10, 5\}$sec is 8, 6, 6 and 10 for the low-relevance category; 12, 11, 14 and 26 for the medium-relevance category; and 39, 64, 89 and 168 for the high-relevance category, respectively. Interestingly, similar ORs' dimensions are observed when considering a different city, with the ORs from Beijing's dataset presenting a median number of cells per OR of 7 for the low-relevance category; 13, 15, 19 and 47 for the medium-relevance category; and 29, 59, 81 and 238 for the high-relevance category, when considering a $T_{thresh} = \{40, 20, 10, 5\}$sec, respectively.

Note that the ORs extracted for $T_{thresh} = 5$sec are the largest in both datasets. Defining small time windows for the OR extraction increases the number of OLCs that can be aggregated, forming larger ORs in the clustering step. Additionally, as high relevance ORs are often personal PoI and indoor, their larger size is due to large indoor GPS errors.

These results show that, *although collected in different cities and time periods, the offloading sites present very similar characteristics evidencing that the laws driving users' mobility of both datasets are similar.*

VI. Identifying Offloading Opportunities

In this section, we inspect offloading opportunities addressing the second question presented in the introduction. Contrarily to the Geolife data, the dataset from Porto provides precise information regarding the start and the end of each

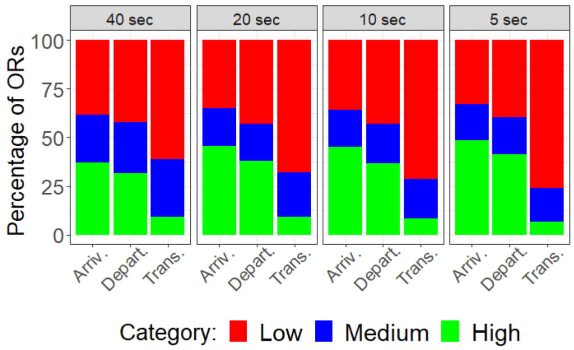


Fig. 8: Relevance of the types of ORs .

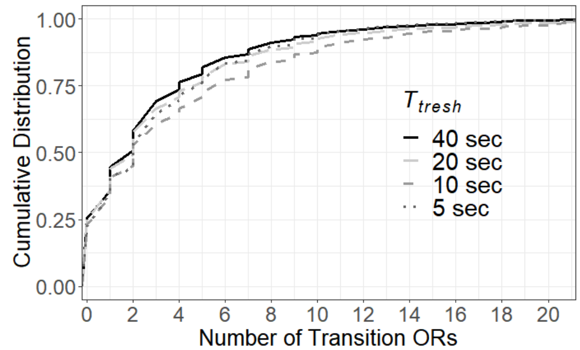


Fig. 9: Number of transition ORs visited per trip.

user's trip, and therefore, it is preferred for this analysis. Thus, we first characterize ORs by type, and then, evaluate the offloading opportunities provided by transition ORs.

A. Type of OR

To understand how ORs are being encountered during a trip, ORs are classified into three types: departure, arrival or transition OR. If the departure location of a user's trip belongs to an OR that OR is classified as a departure OR. In the same way, if the trip's arrival location belongs to an OR, that OR is identified as the arrival OR. All the remaining ORs visited while the user is travelling are identified as transition ORs.

When inspecting the relevance of ORs according to their type, Figure 8 shows that the majority of the departure and arrival ORs exhibit a medium or high relevance, while the majority of transition ORs have low relevance. The reason for having slightly more high-relevance arrival ORs than departure ORs is due to a small delay of the crowdsensing application on starting data collection at the beginning of a trip (Section-III). Hence, for some of the trips, the departure location ended up being associated with a transition OR. Transition ORs correspond to stops during the trip caused by factors such as waiting for public transportation, traffic congestion, traffic lights, etc. These stops have a higher degree of randomness than the stops that correspond to the start and end of trips, contributing to decrease the probability of these ORs to be visited frequently. This will be further analyzed in Section VII.

B. Availability of ORs

To evaluate the offloading opportunities offered to users while travelling, the number of transition ORs visited per trip is evaluated (Figure 9). For 50% of the trips, users visit 1-5 ORs while in transit, revealing that several offloading opportunities exist to be explored apart from the trip destination. When inspecting the impact of considering distinct T_{thresh}, it shows that a T_{thresh} = 40sec provides slightly less transition ORs per trip when compared to smaller T_{thresh}. This is because smaller T_{thresh} allow to extract ORs when users travel at higher speeds as they need to stay less time in a location for it to be considered an OR. In Figure 9, we also observe that for 25% of the trips users can not offload while in transit as these trips do not provide any transition OR. In these cases, the user can only offload at the trip destination and the maximum offloading delay will depend on the trip's duration.

When analyzing if users revisit the same transition OR during the trip, it was verified that this scenario is rare (less that 5% of the ORs are revisited in the same trip). *This suggests that offloading strategies that rely on multiple contacts between a user and APs, e.g. IP caching to reduce connection set-up time [24], would not be effective when the goal is to offload during the trip.* This results are therefore aligned with the previous finding when exploring the interactions between users and APs in urban scenarios.

C. Time Window for Offloading

The time available for offloading represents the time window that mobile devices can use to offload, therefore, it can be estimated based on the time period a user stays inside a transition OR during a trip. Figure 10 shows the users' average sojourn time in transition ORs per trip. As shown, 50% of the users spent more than 66, 56, 56 and 48sec on average inside transition ORs for $T_{thresh} = \{40, 20, 10, 5\}$sec, respectively.

The possibility of offloading not only depends on the time window to offload in each OR but also on the overall coverage provided by ORs during the trip in the time domain. For instance, offloading performance is maximized when users spend most of the trip duration inside ORs. Computing this time parameter, Figure 11 shows that 75% of the users stay, on average, more than 22, 25, 30 and 37% of the trips duration inside ORs for $T_{thresh} = \{40, 20, 10, 5\}$sec, respectively. Smaller T_{thresh} lead to larger ORs (Section V-B), which increases the spatial coverage of ORs and the overall time users are within ORs, and consequently, the ability to offload. These results show that *the users' mobility clearly provides offloading opportunities while users are in transit.*

D. Offloading Delay

Depending on the traffic's requirements, applications can defer data transmission until a WiFi connection is available. This deferred transmission time, here called offloading delay, can be estimated based on the time users take to move between transition ORs. In this case, the observed travel time between transition ORs can serve as a lower bound for the minimum offloading delay regarding the offloading process.

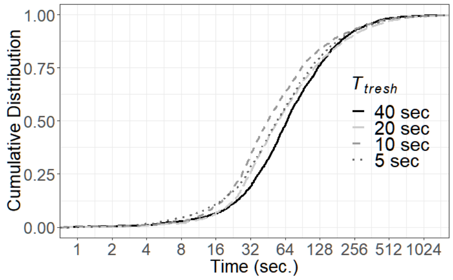


Fig. 10: Users' average sojourn time in transition ORs per trip.

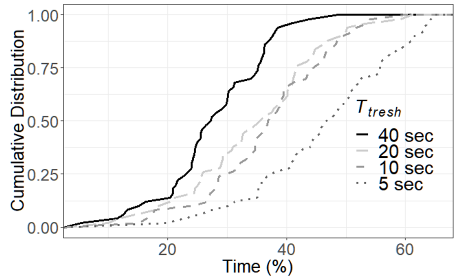


Fig. 11: Users' average percentage of time inside ORs per trip along trips' duration.

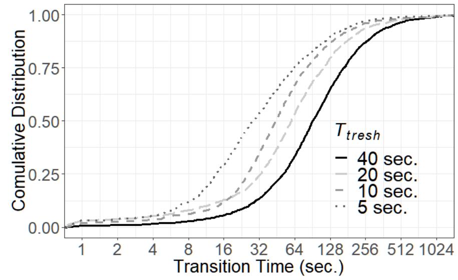


Fig. 12: Users' average travel time between ORs.

Figure 12 shows the users' average travel time between transition ORs. As shown, 50% of the users spend, on average, less than 96, 62, 48 and 28sec travelling between ORs for $T_{thresh} = \{40, 20, 10, 5\}$sec, respectively, evidencing that ORs extracted with lower T_{thresh} values provide lower offloading delays to the users. Smaller T_{thresh} allow higher mobility to users (Eq. 1) creating larger ORs (Figure 7a), which leads to a decrease in the overall offloading delay at the expense of smaller time windows for offloading.

To conclude, the results show that small offloading delays along with high-temporal coverage of ORs confirm a clear opportunity to offload data while the users move in the city.

VII. Next Offloading Region's Prediction

To take advantage of the offloading opportunities provided to users, it is relevant to predict the next OR to be visited by a user. This allows offloading systems to take preemptive actions to manage the offloading process, e.g., use strategies to improve link-setup-time with APs in the ORs [24] or, ultimately, to support the offload decision. Thus, the following study will focus on the spatial OR prediction task bringing the answer to the third question presented in the introduction. Initially, we explore the theoretical predictability of the mobility patterns. Next, we use Markov chains and Machine Learning predictors as attempts to achieve that predictability.

A. Theoretical Predictability

The theoretical predictability of a sequence is correlated with uncertainty, which is usually measured by the entropy rate in information theory. This measurement has been adopted in human mobility studies to establish bounds on predictability under certain assumptions [12], [13], [25]. Here, the three most common entropy measures are assigned to each user mobility pattern represented by the sequence of transitions between its ORs, namely: (i) random entropy $S_u^{rand} = log_2(N_u)$, where N_u is the number of distinct ORs visited by the user u, assuming that each OR is visited with the same probability, ignoring both the frequency of visits and the temporal order of the visits to the ORs; (ii) temporal-uncorrelated entropy $S_u^{unc} = -\sum_{i=1}^{N} p_i log_2 p_i$, obtained by ignoring just the temporal order information of visits to the ORs and applying the entropy formula to the frequency of visits; and, finally, (iii) the true entropy $S_u^{true} = (\frac{1}{n}\sum_i L_i)^{-1} ln(n)$, that considers both the frequency and the temporal order information of the visits to the ORs.

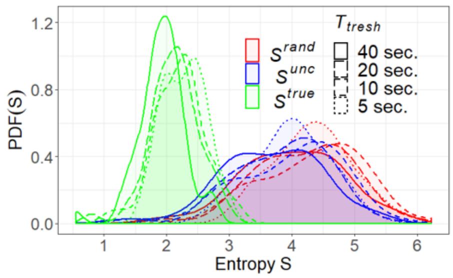
Figure 13 shows the entropy distributions across users from both Beijing and Porto datasets and, naturally, $S^{true} \leq S^{unc} \leq S^{rand}$. Moreover, for both datasets, the results show that considering a small T_{thresh} increases the entropy of the time series. As shown in Figure 13a, for users from the city of Porto, S^{rand} peaks approximately at ≈ 4.4 and ≈ 4.7 for a $T_{thresh} = 40$ and 5sec respectively, indicating that a user who chooses randomly his next OR could be found, on average, in any of $2^{4.4} \approx 21$ or $2^{4.7} \approx 26$ ORs. However, it is important to note that for $T_{thresh} = 40$ and 5sec, S^{true} peaks at ≈ 2 and ≈ 2.5, indicating that the real uncertainty in the users' next OR is now $2^2 \approx 4$ and $2^{2.5} \approx 5.6$ ORs. Interestingly, the same entropy behavior and values can be identified in the users from Beijing, suggesting that the uncertainty on the mobility for offloading is independent of the users' geographical location at a country scale or even at a city scale.

Given the entropy $S^\bullet$ of an individual who moves between N ORs, the upper bound of mobility predictability is given by the probability $\Pi^\bullet$, which represents the maximum accuracy that can be achieved by a prediction algorithm. This probability $\Pi^\bullet$ is expressed by $S^\bullet = H(\Pi^\bullet) + (1 - \Pi^\bullet)log_2(N-1)$ with the binary entropy function $H(\Pi\bullet) = -\Pi^\bullet log_2(\Pi^\bullet) - (1 - \Pi^\bullet)log_2(1-\Pi^\bullet)$. This quantity is subject to Fano's inequality (we refer the reader to [12], [26] for more details). After determining $\Pi^\bullet$, and as Figure 14 illustrates, for both cities, the Π^{true} peaks between ≈62% and ≈70% for all T_{thresh}, a value considerably low when compared to related work. For instance, a Π^{true} of 93% is obtained when considering human mobility between personal PoIs [12]. *This result clearly shows that the prediction task for offloading is more challenging, though possible, motivating the need for further investigation of the mobility properties under offloading conditions.*

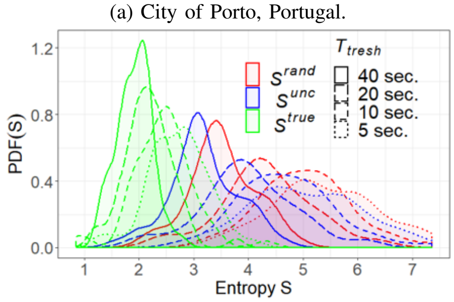
B. Markov Chain Predictor

Markov Chain (MC) is the most commonly used predictor for users mobility [17], [18]. To evaluate the user's mobility between ORs, we model the mobility behavior as a discrete stochastic process using a MC consisting of a set of states $S = \{S_1, ..., S_i, ...S_n\}$, where S_i corresponds to OR i visited by the user, and a transition probability matrix P, where each element p_{S_i,S_j} represents the user probability of moving from S_i to S_j, with $(i, j) = \{x \in \mathbb{N} | 1 \leq x \leq n\}$ and $i \neq j$.

For each user and day period, the sequence of ORs visited is extracted in a chronological order. The first 75% of the sequence is used as a training set to build the transition

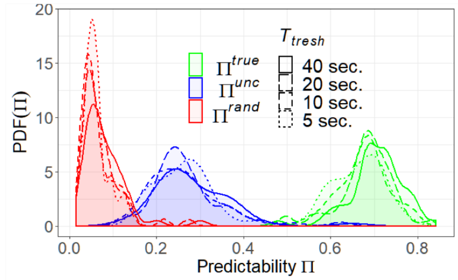


(a) City of Porto, Portugal.

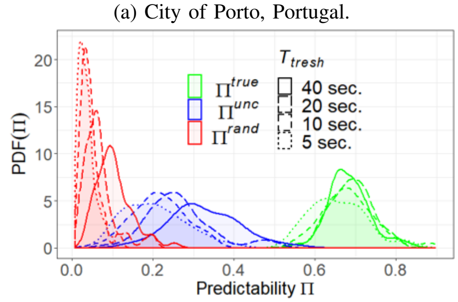


(b) City of Beijing, China.

Fig. 13: Entropy values for the users.

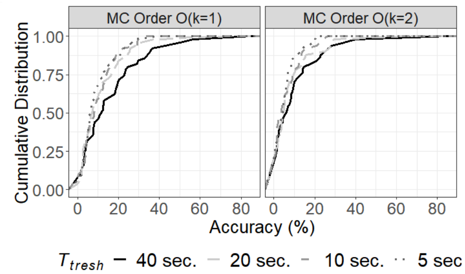


(a) City of Porto, Portugal.

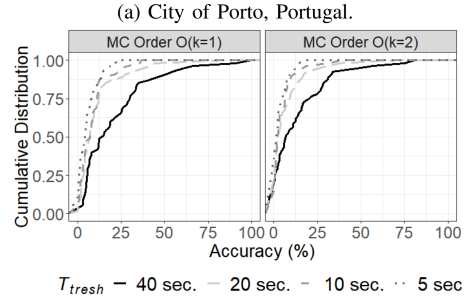


(b) City of Beijing, China.

Fig. 14: Predictability values for the users.



(a) City of Porto, Portugal.



(b) City of Beijing, China.

Fig. 15: MC predictor' accuracy for all users.

probability matrix P and the remaining 25% as the testing set to validate the predictions. Only users with a sequence of visited ORs larger than 10 were considered. For each prediction, the current state S_i is defined as the last OR of the training sequence and the prediction of the next OR to be visited S_j is performed selecting the OR with higher transition probability p_{S_i,S_j}. In the case of two ORs with the same transition probability, one is chosen randomly. If there is no information regarding the current OR in P (e.g., the first time the user visits this OR), it is assumed that the predictor fails as it cannot predict the next OR. The predictor evaluates in runtime – after each prediction, the training sequence is updated with the right prediction, and a new P is computed.

The accuracy of the predictor is evaluated as the number of correct predictions over the total number of predictions. As shown in Figure 15, the MC predictor presents similar results for both cities. An accuracy below 23% and 13% is achieved for 75% of the users from the city of Porto (refer to Figure 15a), when considering only the current location (k=1) and respectively, a $T_{thresh} = 40sec$ and $T_{thresh} = \{20, 10, 5\}sec$. Although an increase of $\approx$ 10% in the MC predictor can be seen in the users from Beijing for the $T_{thresh} = 40sec$ (refer to Figure 15a), these values are considerably smaller that the upper bound presented in Figure 14. This accuracy is even worse when past history is considered, namely, the current and the previous OR (k=2). The reason is that Markovian predictors have a large probability space that increases quickly following a power law with order k. Therefore, when $k > 1$, these predictors may suffer from insufficient samples.

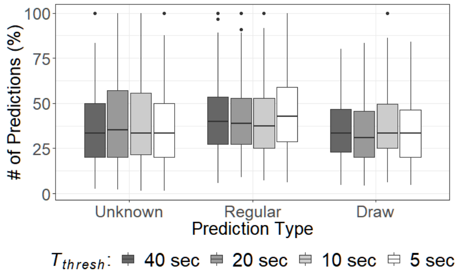
Attending to such poor performance when compared to the theoretical upper bound, we first investigate the possible causes leading to the MC predictor low accuracy and then propose and explore approaches to improve its performance.

*1) **Inspecting MC Predictor**:* A matrix P built using a short history (length of the sequence of transitions between ORs) may not capture the diversity of users routines (movement between ORs). To test this hypothesis, the Pearson correlation coefficient (ρ) between the accuracy of the predictions and the sequences length was computed, and no correlation was found indicating that small sequences may not be responsible for the predictor's poor performance. Instead, a negative correlation was found between the predictor's accuracy and the number of unique ORs of the sequence, $\rho = \{-0.38, -0.29, -0.30, -0.29\}$ for $T_{thresh} = \{40, 20, 10, 5\}$sec, respectively. Similar (ρ) values were found using Beijing's dataset, $\rho = \{-0.46, -0.33, -0.31, -0.27\}$ for $T_{thresh} = \{40, 20, 10, 5\}$sec respectively. This evinces that novelty, which consists on finding new ORs may have more impact on the predictors' performance than the sequence length.

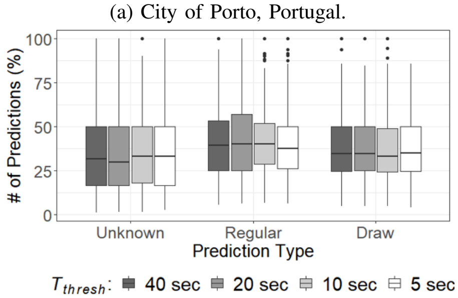
The appearance of new ORs in a sequence will make the predictors fail the prediction in two scenarios, namely, when: i) the current state is a new OR, since there is no information in matrix P and, therefore, the prediction fails; and ii) the next state is a new OR since the predictor can only predict what is in P (what has occured before). To further study the first scenario, we define three types of prediction based on P information during the decision phase, namely:

- *Unknown-based prediction:* occurs when there is no information in P regarding the current state (new OR);
- *Draw-based prediction:* occurs when two or more ORs in P have the same transition probability;
- *Regular-based prediction:* occurs when only one OR exhibits the highest transition probability.

Figure 16 shows the distribution of predictions for each prediction type. It is clear that in both datasets the number of predictions is evenly distributed among the prediction types, showing that the majority of predictions are not regular-based.

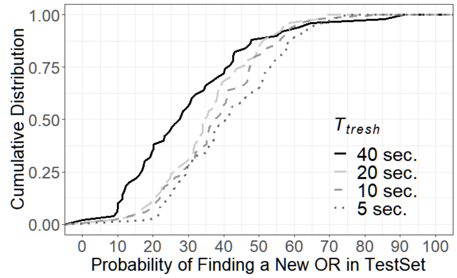


(a) City of Porto, Portugal.

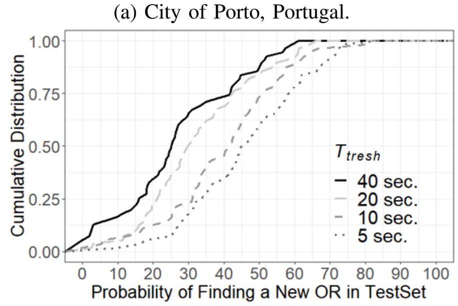


(b) City of Beijing, China.

Fig. 16: Prediction type distribution.

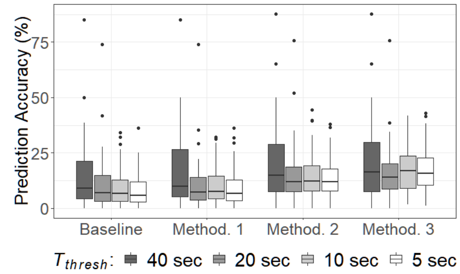


(a) City of Porto, Portugal.

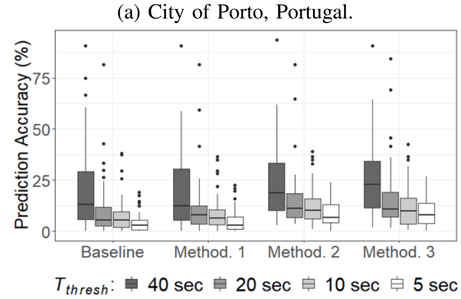


(b) City of Beijing, China.

Fig. 17: Users' probability of finding a new OR.



(a) City of Porto, Portugal.



(b) City of Beijing, China.

Fig. 18: Improved MC predictor's accuracy.

In addition, unknown-based predictions, which are nearly one-third of predictions, are bound to fail.

To further inspect the second scenario mentioned above, we compute the probability of the next OR being a new OR for each sequence. For that, an OR is considered a "new OR" if it appears in the sequence for the first time, or a "repeated OR" otherwise. Note that the predictor is an online predictor, consequently, even after training, P is updated after each prediction. To capture this behavior, the probability of finding a new OR is given by the fraction of "new ORs" in the test sequence over the test sequence length, see results in Figure 17. For the sequences from Porto's dataset in Figure 17a, the median probability of the next OR being a "new OR" is 27% for T_{thresh} = 40sec. This means that an online MC predictor will fail, on average, 27% of its regular-based predictions. This value increases to ≈40% for a T_{thresh} = 5sec values. Similar results can be observed in Beijing's dataset in Figure 17b. Smaller T_{thresh} values increase the probability of finding new ORs as the spatial-temporal constraint (see Section IV-A) for a location to be considered an OR is relaxed, increasing the number of unique ORs extracted from a mobility trace.

*2) **Improving MC Predictor**:* Since most of the prediction types are draw-based and unknown-based, we combine different strategies to improve the MC predictor's performance in these cases, evaluating also their impact. When the MC predictor faces a draw-based prediction the following strategies are applied to the set of ORs with the same transition probability in P: i) randomly select one of the ORs as next OR; or ii) select the closest OR to the current OR. In the other side, when the MC predictor faces an unknown-based prediction the following strategies are applied to select the next OR considering all ORs present in P: i) no OR is selected and the prediction fails; ii) select the closest OR to the current OR; or iii) select the most visited OR in P. While the distance-based strategies try to capture the human mobility tendency to favor shorter paths, history-based strategies such as, *Most Visited OR*, try to capture the regularity behavior of human mobility. Table III shows the different combinations of these strategies when evaluating the MC predictor performance.

TABLE III: Strategies to improve MC predictor.

	Prediction Type	
Strategies	**Draw**	**Unknown**
Baseline	Random	None
Method 1	Closest OR from P	None
Method 2	Closest OR from P	Closest OR
Method 3	Closest OR from P	Most Visited OR

Figure 18 shows the MC predictor accuracy for both dataset for a k = 1 when applying the strategies described above. Solving the ties by selecting the closest OR (Method 1) has a minor impact on the predictor's performance in both datasets. However, solving the unknown-based predictions improves the accuracy around two times in both datasets. Similar results were obtained for methods 2 and 3 revealing no difference between distance-based and historic-based strategies on the unknown-based predictions.

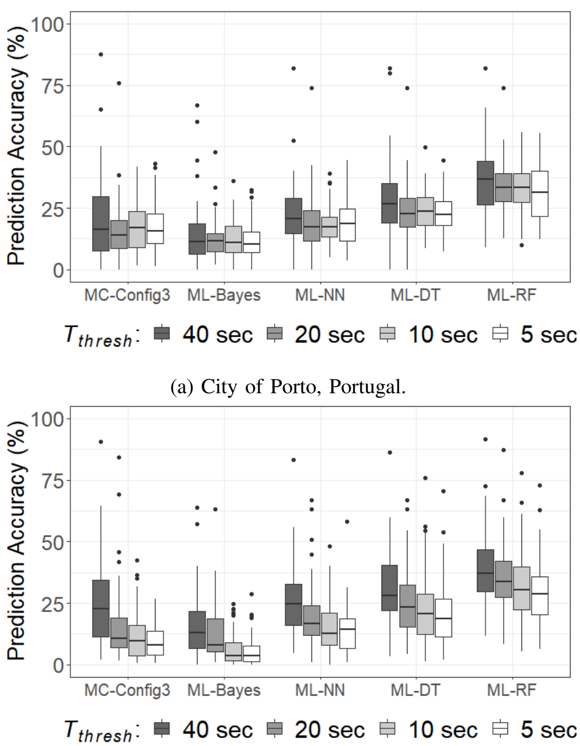
Even with the improvements in the predictors performance, their median accuracy is less than 24% for both Porto and Beijing datasets. *Although Markovian predictors have been proposed as feasible solutions to predict mobility between users' PoI [17], [18], the results show that their performance is significantly reduced in offloading scenarios.*

C. Machine Learning Predictors

From the previous analysis, it was clear that the MC predictor is not capable of predicting the next OR conveniently. Other types of predictors are Machine Learning (ML) predictors. In our mobility prediction task, the input consists of a series of

TABLE IV: Mobility features for predicting the next OR.

Features Domain	Feature	Type	Description
Location	Current OR (COR)	Nominal	OR where the user is
	Previous OR (POR)	Nominal	OR from where the user comes from
	Cell in COR	Numeric	Coordinates of the grid cell (11x11m square) where the user is inside the OR
Time	Duration in COR	Nominal	Time spent in the current OR
	Day	Nominal	Weekday name
	Hour	Numeric	Hour of the day
Relevance	Category	Nominal	Relevance Category of the COR



(a) City of Porto, Portugal.

(b) City of Beijing, China.

Fig. 19: ML predictors' accuracy.

contextual features associated with the current OR, while the output is the next OR. Context features regarding the current OR were divided into three feature domains: location, time, and relevance. Table IV details the features extracted from the users' mobility traces. The nominal variables were converted to binary features using one-hot encoding. Then, different types of ML predictors were used, namely, decision trees (DT), neural networks with a single hidden layer (NN), bayesian (Bayes) and random forests (RF) using the CARET package in R with the algorithms C5.0, avgNN, bayesglm and ranger, respectively, and the default parameter settings.

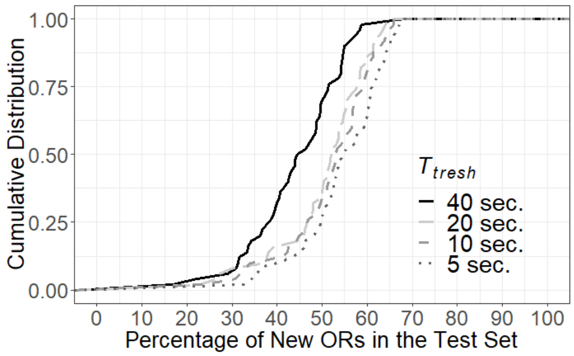
Figure 19 shows the accuracy of ML predictors using a cross-validation of $k = 3$. With the exception of the Bayesian predictor, the ML predictors perform better than the MC predictor, with RF presenting the best results for both datasets with median accuracy between $\approx$ 30% and $\approx$ 37% for all T_{thresh} values. However, it is important to note that considering higher T_{thresh} improves the prediction accuracy, being this more evident in the Beijing dataset, see Figure 19b. Note that the models considered could be further fine-tuned or even replaced by others which could lead to better results. However, our goal is not to produce a state-of-the-art prediction model, but instead to inspect the limitations and sources of predictability in an offloading scenario.

Contrary to the MC predictor, the online version of the ML predictors is not considered as the training phase may have a high-processing cost, being likely unfeasible when the offloading process is handled in mobile devices. Therefore, new ORs can have more impact on the ML predictors performance as the predictors can only predict ORs that are present in the initial training sequence (75% of the sequence). Figure 20 shows the percentage of new ORs in the testing sequence when compared to the training sequence. Once more, similar results can be observed for both datasets where the lower number of new ORs in the test set is obtained for a T_{thresh} = 40sec. Here, the median percentage of new ORs in the test set per user is $\approx$ 45% for T_{thresh} = 40sec, showing that, for half of the users, the maximum accuracy that ML predictors would be able to achieve is $\approx$ 55%. Note that while the users from Porto present a similar number of new ORs in the test set for the other T_{thresh} values, in Beijing, the number of ORs increases as smaller T_{thresh} values are considered. This causes the reduction in the ML predictors performance observed in Figure 19b as higher percentages of new ORs in the test set lead to a decrease in the predictors performance.

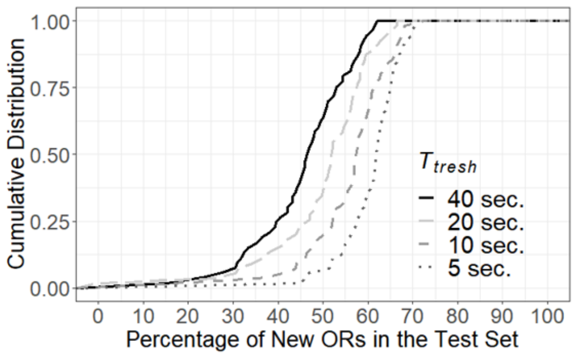
Even though the new ORs phenomenon is more challenging for ML predictors without having an online nature (as the MC predictor used before), the performance of predicting the next OR is higher. *In an offloading scenario, the smaller spatial-temporal constraint determined by T_{thresh} causes the users to discover new ORs constantly, which may count with just a few visits. Even comparing to an online MC predictor, our results show that ML present better performance in predicting offloading opportunities, revealing the importance of contextual information (Table IV) in offloading scenarios.*

VIII. Mobility Properties for Offloading

The previous section shows that looking at mobility from an offloading perspective imposes different assumptions when compared to human mobility as studied in related work. Thus, previous conclusions on human mobility do not apply, e.g. high performance of MC predictors [17], [18]. In this section, we aim to shorten this gap by inspecting the properties of human mobility in an offloading scenario and their impact on designing offloading systems finally answering the fourth and last question presented in the introduction.



(a) City of Porto, Portugal.



(b) City of Beijing, China.

Fig. 20: New ORs in the test set when compared to the training set.

A. Exploration Phase Characteristics

We define the *exploration phase* of a user as the time to visit each OR for the first time. As discussed in Section VII, the high probability of a user to find new ORs is the main challenge to the user's mobility predictability. One may argue that this occurs due to the lack of historical information, and that the new OR phenomenon will have less impact if longer mobility traces are considered.

To evaluate this hypothesis, we compute and analyse the users' discovery factor DF. The discovery factor DF_i^u of a user u in the i^{th} day of its mobility trace is defined as the percentage of ORs visited by that user u at the i^{th} day, with respect to the total number of different ORs visited by the same user during D days. We compute the DF during the commuting periods for users with 15 days of historical data.

Figure 21 shows the users' discovery factor for the morning commuting period (DP2); similar results were found for DP4. *The similarity of the results from both datasets shows a clearly indication that the user's exploration phase on an offloading scenario is common even across datasets collected from different countries at different times.* When considering only medium and high-relevance ORs, it can be observed that the users' exploration phase is very short - 50% of users visit all their medium and high-relevance ORs after just 3 days (median DF_3^u = 100%). However, when considering also the low-relevance ORs (then considering all ORs) it can be seen that users are constantly discovering new low-relevance ORs every day. This can be seen by the tendency line of the discovery factor indicator which increases almost linearly with the number of days. As shown in Figure 6, most low-relevance ORs are rarely visited by the users (max. distribution peaking approximately at 7%) which makes them very likely to correspond to random stops taken by the users while moving, and therefore, very probable to continue to appear even if a larger historical data is considered. *This shows that the new OR phenomenon is due to the mobility characteristics in an offloading scenario and not due to the lack of historical data: In [27], [28], authors discuss the hardness in prediction given by the novelty component in mobility. Therefore, considering longer learning periods for offloading systems is unlike to improve their capacity of predicting mobility.*

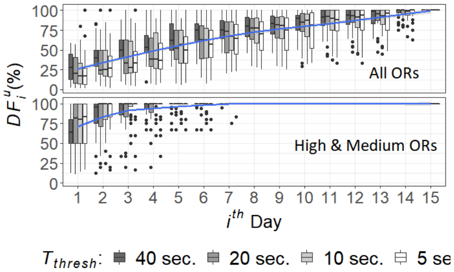
B. Mobility Regularity Effects

The regularity of users' mobility is characterized by the ORs' visiting patterns, which might be used to improve mobility predictability. We further evaluate the regularity impact by answering the question: Can specific characteristics of regularity be used to improve the mobility predictability between ORs?

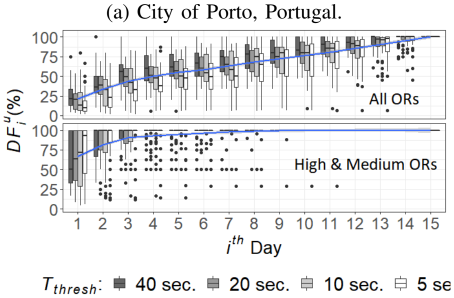
*1) **Relevance Filter**:* Offloading systems may try to improve the user's mobility regularity using relevance as a filter, e.g. considering only ORs that have a specific level of relevance to the users.

ORs with a relevance lower than 7%, 10% and 15% were filtered from the sequence of visited ORs and the predictors re-trained. Figure 22 shows the accuracy of the Random Forest predictor (the best from Figure 19) for the different relevance filters, for both datasets. The results show that the predictor's accuracy improves when ORs with low relevance are removed. In fact, the prediction accuracy improves approximately 20% in both datasets when a 15% relevance filter is used, reaching a median accuracy of approximately 55% for T_{thresh} = $\{40, 20, 10\}$ and 60% for T_{thresh} = 5sec. Higher values for the relevance filter could be used to improve the mobility predictability in an offloading scenario, however, this affects the offloading opportunities. This trade-off will be further studied in Section VIII-C.

*2) **Multiple OR Prediction**:* Depending on how the offloading process is managed, offloading systems may consider more than one OR in advance to prepare the offloading process, e.g., launching mechanisms to provide seamless network access to mobile devices. In this case, all the setup actions anticipating the users' arrival can be performed in two or more ORs, and the final OR choice can be made online using contextual information such as the current user's path. This scenario is only convenient to offloading systems if the uncertainty associated with the prediction of the next OR is concentrated in a small subset of ORs, for each prediction. Thus, the problem formulation for the prediction task can be loosened: what is the ML predictor accuracy if we consider the N most probable ORs as being the next OR? For this, in each prediction, the top N ORs with the highest probability of being the next OR visited by the user are selected. Then, a prediction is assumed to be correct if one of the top N selected ORs is, in fact, the next OR visited by the user. For this, the RF predictor is used and only ORs with relevance greater than 15% are considered.

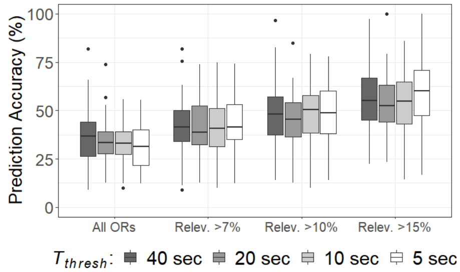


(a) City of Porto, Portugal.

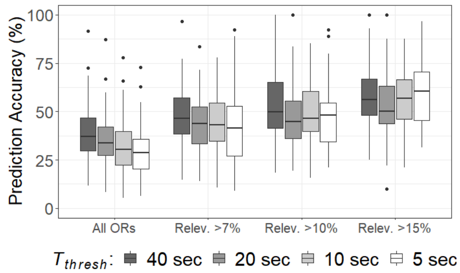


(b) City of Beijing, China.

Fig. 21: User's discovery factor for the DP2 day period.

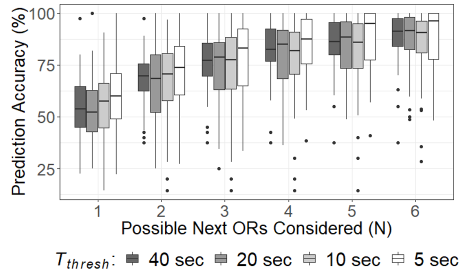


(a) City of Porto, Portugal.

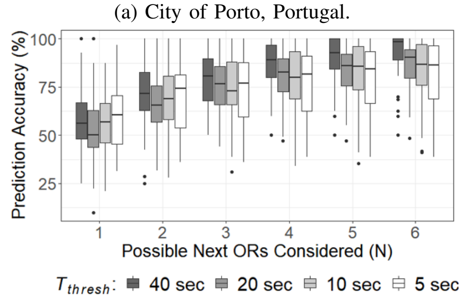


(b) City of Beijing, China.

Fig. 22: Random Forest predictor's accuracy for different relevance filters.



(a) City of Porto, Portugal.



(b) City of Beijing, China.

Fig. 23: Random Forest predictor's accuracy for N possible next ORs.

Figure 23 shows the predictor's accuracy when considering that the next OR will be one of the top N ORs selected by the predictor, where $N = \{x \in \mathbb{N} | 1 \leq x \leq 6\}$. As shown, for both datasets, the predictors' accuracy increases significantly when considering multiple options as the next OR; in these case, from $\approx 15\%$ for $N = 2$ to $\approx 27\%$ for $N = 4$. *This reveals that the uncertainty of predicting the next visited OR is concentrated in a small subset of ORs which can be leverage by offloading systems. However, considering more than four possible next ORs does not bring significant added value as after $N > 4$ only marginal improvements are obtained.*

C. Mobility Predictability Trade-offs

The deterministic or opportunistic nature of the offloading strategies used by offloading systems depends on the users' mobility predictability. Ideally, high predictability is preferable as it allows using more sophisticated and deterministic strategies anticipating the users arrival to an OR to maximize offloading performance. In the other hand, opportunistic strategies should be preferred when the predictability is lower. In this case, the offloading process can initiate opportunistically every time a user enters an OR without any anticipatory action needed to be taken.

As demonstrated in the previous sections, despite mobility predictability being a challenge in an offloading scenario, it can be improved if: (i) more than one OR is considered as the next OR; and (ii) low-relevant ORs are not considered during the prediction process. However, most of the transition ORs are low-relevance ORs and removing them may affect the offloading opportunities provided to users in transit. Figure 24 and Figure 25 show the impact of removing low-relevance ORs on the number of transition ORs and the percentage of time inside ORs per trip (Porto's dataset), respectively. For clarity purposes, only results for $T_{thresh} = 10sec$ are shown as similar results were found for other T_{thresh}. Removing low-relevance ORs decreases the offloading opportunities while the user is travelling. In fact, when considering only ORs with relevance higher than 15%, the number of trips where users cannot offload while travelling increases from 25% to 38% (Figure 24). Moreover, when offloading while travelling is possible, the median percentage of time spent by users inside ORs decreases from 37% to 28% (Figure 24). *These results clearly show a trade-off between the mobility predictability and the offloading opportunities. Therefore, to maximize the offloading opportunities to users in transit, an efficient offloading strategy should combine opportunistic and predictive strategies.*

D. Offloading Sites Considerations

In a mobile data offloading scenario WiFi APs should provide Internet connectivity in the offloading sites. Larger ORs may require more APs and consequently more handovers to provide full WiFi connectivity during the offloading process. Figure 26 shows the maximum coverage of the offloading sites computed as the maximum euclidean distance between two geographical coordinates belonging to the same OR. Similar to the number of cells (see Section V-B), the maximum coverage of the ORs is very similar in both datasets. The horizontal lines in Figure 26 represent the typical coverage of an AP (44m) obtained in [19] through WiFi measurements in an urban scenario. *As the vast majority of transition ORs belong to the low-relevance category, this shows that mobile devices will not need to perform multiple handovers during the offloading process as few APs, e.g. two APs, can provide full WiFi coverage to the transition ORs.*

Attending that most APs in [19] were were inside buildings and homes, *this indicates that already deployed WiFi infrastructure can be leveraged by offloading systems to offload since*

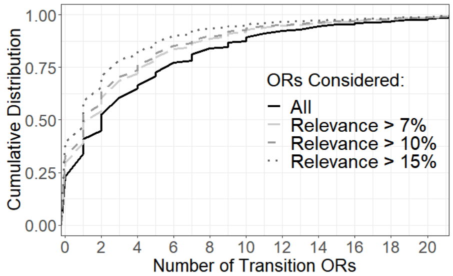


Fig. 24: Number of transition ORs per trip (T_{thresh} = 10 sec).

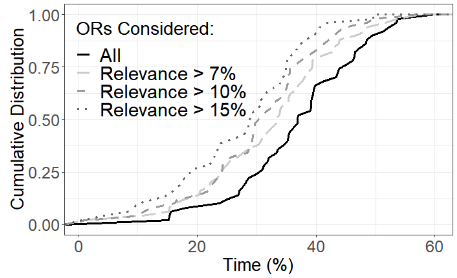


Fig. 25: Users' average time inside ORs per trip (T_{thresh} = 10) sec.

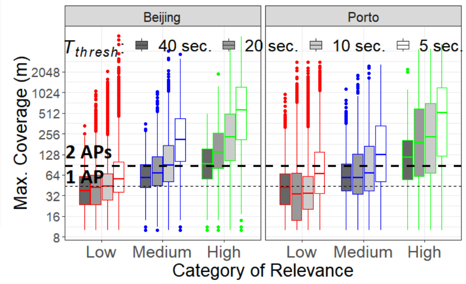


Fig. 26: Maximum coverage of the ORs.

they provide a feasible coverage. Therefore, large studies such as [29]–[31] that characterize WiFi APs deployments at city scale can be used to design offloading decision models that can take advantage of the already deployed WiFi infrastructure.

IX. Conclusion

This work proposes the use of granular human mobility profiles for informing offloading strategies during commuter trips. ORs can be extracted from individual trajectories using unsupervised learning methods, that were validated on data collected from two different cities at different times. Offloading opportunities were investigated and the results show that users can offload while in transit considering the ORs' availability along with reasonable sojourn times and small offloading delays between OR.

Mobility predictability for offloading in commuter trips revealed to be much more challenging when compared to mobility predictability between PoI. Here, ML predictors outperform MC predictors ($\approx$ 37% vs. $\approx$ 12% acc.), revealing the importance of a contextual feature when predicting the next OR to be visited by a user. The results indicate that the users' high probability of finding new ORs is the main cause for predictors' poor performance. However, inspecting the users' exploration phase showed that this behavior is due to the mobility characteristics in offloading scenarios, and considering longer learning periods will not improve mobility predictability. Nevertheless, the results also show that a significant improvement in mobility predictability can be obtained if mobility regularity properties are used. Here, offloading systems can consider up to four possible ORs as the next ORs and use a relevance filter to improve prediction performance. This implies the implementation of both opportunistic and deterministic strategies to maximize offloading performance. Our results also shown that mobile devices will not need to perform multiple handovers inside an OR during the offloading process as ORs can be covered by a small number of APs.

Finally, the findings can be considered general since they are based on two datasets collected independently on two sides of the world. A high degree of similarity between the results from both Porto and Beijing datasets was found, indicating that offloading mobility is highly dependent on how regular people move in cities. Therefore, the findings and conclusion of this work are likely to be generalized to other urban scenarios.

This work opens up new paths of research on the design of personalised offloading systems that adapt to the user's mobility. As next steps we see the assessment of different types of offloading mobility profiles, e.g. considering transportation mode, and the evaluation of the offloading opportunities and predictability for each. The design and evaluation of such systems considering individual mobility and specific applications is another line of future work.

References

[1] C. V. N. Index, "Cisco visual networking index: Global mobile data traffic forecast update, 2017–2022 white paper," 2017, (accessed 14-01-2020). [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-738429.html

[2] M. H. Cheung and J. Huang, "Dawn: Delay-aware wi-fi offloading and network selection," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 6, pp. 1214–1223, 2015.

[3] H. Zhou, H. Wang, X. Li, and V. C. Leung, "A survey on mobile data offloading technologies," *IEEE Access*, vol. 6, pp. 5101–5111, 2018.

[4] K. Lee, J. Lee, Y. Yi, I. Rhee, and S. Chong, "Mobile data offloading: How much can wifi deliver?" *IEEE/ACM Transactions on networking*, vol. 21, no. 2, pp. 536–550, 2012.

[5] Z. Lu, X. Sun, and T. La Porta, "Cooperative data offload in opportunistic networks: From mobile devices to infrastructure," *IEEE/ACM Transactions on Networking*, vol. 25, no. 6, pp. 3382–3395, 2017.

[6] H. Zhou, J. Chen, H. Zheng, and J. Wu, "Energy efficiency and contact opportunities tradeoff in opportunistic mobile networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 5, pp. 3723–3734, 2015.

[7] Y. Im, C. Joe-Wong, S. Ha, S. Sen, M. Chiang *et al.*, "Amuse: Empowering users for cost-aware offloading with throughput-delay tradeoffs," *IEEE Transactions on Mobile Computing*, vol. 15, no. 5, pp. 1062–1076, 2015.

[8] F. Mehmeti and T. Spyropoulos, "Performance modeling, analysis, and optimization of delayed mobile data offloading for mobile users," *IEEE/Acm Transactions on Networking*, vol. 25, no. 1, pp. 550–564, 2016.

[9] D. Ciullo, T. Spyropoulos, N. Nikaein, B. Jechoux, and G. Sarantidis, "Sizing up user traffic: Flow-based mobile data offloading over wifi," in *2019 IEEE 20th International Symposium on" A World of Wireless, Mobile and Multimedia Networks"(WoWMoM)*. IEEE, 2019, pp. 1–9.

[10] Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma, "Mining interesting locations and travel sequences from gps trajectories," in *Proceedings of the 18th international conference on World wide web*. ACM, 2009, pp. 791–800.

[11] X. Cao, G. Cong, and C. S. Jensen, "Mining significant semantic locations from gps data," *Proceedings of the VLDB Endowment*, vol. 3, no. 1-2, pp. 1009–1020, 2010.

[12] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.

[13] G. Smith, R. Wieser, J. Goulding, and D. Barrack, "A refined limit on the predictability of human mobility," in *2014 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 2014, pp. 88–94.

[14] Q. Liu, S. Wu, L. Wang, and T. Tan, "Predicting the next location: A recurrent model with spatial and temporal contexts," in *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.

[15] T. M. T. Do, O. Dousse, M. Miettinen, and D. Gatica-Perez, "A probabilistic kernel method for human mobility prediction with smartphones," *Pervasive and Mobile Computing*, vol. 20, pp. 13–28, 2015.

[16] Z. Yang, J. Hu, Y. Shu, P. Cheng, J. Chen, and T. Moscibroda, "Mobility modeling and prediction in bike-sharing systems," in *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 2016, pp. 165–178.

[17] X. Lu, E. Wetter, N. Bharti, A. J. Tatem, and L. Bengtsson, "Approaching the limit of predictability in human mobility," *Scientific reports*, vol. 3, p. 2923, 2013.

[18] C. Krumme, A. Llorente, M. Cebrian, E. Moro *et al.*, "The predictability of consumer visitation patterns," *Scientific reports*, vol. 3, p. 1645, 2013.

[19] E. Lima, A. Aguiar, P. Carvalho, and A. C. Viana, "Impacts of human mobility in mobile data offloading," in *ACM CHANTS*, 2018.

[20] J. G. Rodrigues, A. Aguiar, and C. Queirós, "Opportunistic mobile crowdsensing for gathering mobility information: Lessons learned," in *Intelligent Transportation Systems (ITSC), 2016 IEEE 19th International Conference on*. IEEE, 2016, pp. 1654–1660.

[21] Y. Zheng, H. Fu, X. Xie, W.-Y. Ma, and Q. Li, *Geolife GPS trajectory dataset - User Guide*, geolife gps trajectories 1.1 ed., July 2011, geolife GPS trajectories 1.1. [Online]. Available: https://www.microsoft.com/en-us/research/publication/geolife-gps-trajectory-dataset-user-guide/

[22] S. Firdaus and M. A. Uddin, "A survey on clustering algorithms and complexity analysis," *International Journal of Computer Science Issues (IJCSI)*, vol. 12, no. 2, p. 62, 2015.

[23] T. M. Kodinariya and P. R. Makwana, "Review on determining number of cluster in k-means clustering," *International Journal*, vol. 1, no. 6, pp. 90–95, 2013.

[24] C. Pei, Z. Wang, Y. Zhao, Z. Wang, Y. Meng, D. Pei, Y. Peng, W. Tang, and X. Qu, "Why it takes so long to connect to a wifi access point," in *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*. IEEE, 2017, pp. 1–9.

[25] D. D. C. Teixeira, A. Carneiro Viana, M. S. Alvim, and J. M. Almeida, "Deciphering Predictability Limits in Human Mobility," in *ACM SIGSPATIAL*, Nov. 2019.

[26] M. Feder and N. Merhav, "Relations between entropy and error probability," *IEEE Transactions on Information Theory*, vol. 40, no. 1, pp. 259–266, 1994.

[27] L. Amichi, A. Carneiro Viana, M. Crovella, and A. A. F. Loureiro, "From movement purpose to perceptive spatial mobility prediction," in *ACM SIGSPATIAL 2021*, Beijing, China, Nov. 2021. [Online]. Available: https://hal.inria.fr/hal-03444658

[28] D. Teixeira, J. Almeida, and A. Carneiro Viana, "On estimating the predictability of human mobility: the role of routine," *EPJ Data Science*, vol. 10, no. 1.

[29] L. Zhang, L. Zhao, Z. Wang, and J. Liu, "Wifi networks in metropolises: From access point and user perspectives," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 42–48, 2017.

[30] G. Capdehourat, F. Larroca, and G. Morales, "A nation-wide wi-fi rssi dataset: Statistical analysis and resulting insights," in *2020 IFIP Networking Conference (Networking)*. IEEE, 2020, pp. 370–378.

[31] E. Lima, A. Aguiar, and P. Carvalho, "Offloading surrogates characterization via mobile crowdsensing," in *Proceedings of the First ACM Workshop on Mobile Crowdsensing Systems and Applications*. ACM, 2017, pp. 7–12.

Emanuel Lima received his B.Sc's Degree in Electronics and Computers' Networks in 2012 and his M.Sc. in Networks and Communications Services from the University of Minho, Portugal, in 2014. He is currently a Ph.D. candidate at University of Porto (FEUP), Portugal where he develops his work at Instituto de Telecomunicações (IT) Porto and was awarded a Doctoral Scholarship from the Portuguese Foundation for Science and Technology (FCT) in 2015. His research interests include wireless communications, mobile and edge computing focusing on the development of intelligent solutions to improve users' Quality of Experience.

Ana Aguiar is Assistant Professor at the University of Porto (UP) and researcher at Instituto de Telecomunicações (IT), where she coordinates the Network and Information Processing Group at IT Porto. Her research focuses on wireless and mobile systems, concretely connectivity for mobile nodes, mobile IoT, and data science for future mobility and adaptation to climate change. She is adjunct coordinator of the Center of Competences for Future Cities at UP, where she operated the UrbanSense environmental sensors and the crowdsensing tool SenseMyCity. She led the VOCE project on voice stress detection, and participated in more than 15 projects, in 9 of them as co-PI or WP leader. She published more than 70 journal and conference articles in renowned venues, is reviewer for several IEEE transactions journals, and has been part of the organising committee of several top conferences in her field of work (e.g. ACM Mobicom, IEEE WoWMoM, IEEE VNC).

Paulo Carvalho graduated in Systems and Informatics Engineering at University of Minho, Portugal, in 1991, and received his PhD degree in Computer Science from the University of Kent at Canterbury, United Kingdom, in 1997. He is currently Associate Professor in Computer Communications, Department of Informatics, School of Engineering, University of Minho, Portugal. He is actively involved in several postgraduate Scientific Coordinating Committees both at Doctoral and Master level, namely, the Doctoral Programme in Computer Science (MAP-i), a triple-degree joint PhD programme, involving major universities – Minho, Aveiro, Porto; the Master Degree in Informatics Engineering (both Integrated and 2nd cycle); and the Master Degree on Networks Engineering and Telematic Services. He is a member of Algoritmi research center, integrating the Computer Communications and Pervasive Media (CCPM) Group. His main research interests include broadband technologies, multiservice networks, mobile networks, and teletraffic analysis, topics in which he has been publishing actively over the last decades. He is also a TPC member of major international conferences, and a regular reviewer of IEEE/ACM journals.

Aline Carneiro Viana is a Senior Reseach Director (DR) at Inria, where she leads the team TRiBE team. After a 2-years working at the TKN/TU-Berlin in 2009, Germany, she got in 2011 the *habilitation* degree (HDR) from UPMC - Sorbonne Universités. Dr. Viana got her PhD in Computer Science from the UPMC - Sorbonne Universités in 2005. Her research addresses the design of solutions applied to Tactile Internet, perceptive-smart networking and mobile self-organizing networks with the focus on human behavior analytics. She is a recipient of the French Scientific Excellence award (2015-2022), got the Medal of Honor 2021 from the Federal UFG (Brésil) for her relevant society contributions, and was nominated in 2016 as one of the "10 women in networking/communications that you should Watch" (1st-year nomination of N2Women community). She is author of more than one hundred papers, including top-tier conferences and important peer-reviewed journals. She has been involved in the organizing committee of major conferences (e.g. ACM SenSys and Mobicom, IEEE Infocom and LCN) as well in the management of EU, international, and French projects as well the EMBRACE Inria Associate Team (between Inria and 3 Brazilian partners).