Title	Virtual Location-Based Services: Merging the Physical and Virtual World
Author(s)	Von Der Weth, Christian; Hegde, Vinod; Hauswirth, Manfred
Publication Date	2014
Publication Information	Chrsitian Von Der Weth, Vinod Hegde, Manfred Hauswirth (2014) Virtual Location-Based Services: Merging the Physical and Virtual World 21st IEEE International Conference on Web Services
Item record	http://hdl.handle.net/10379/4440

Virtual Location-Based Services:
Merging the Physical and Virtual World

Christian von der Weth   Vinod Hegde   Manfred Hauswirth
Digital Enterprise Research Institute (DERI) – National University of Ireland, Galway (NUIG)
{christian.vonderweth|vinod.hegde|manfred.hauswirth}@deri.org

***Abstract*—Location-based services gained much popularity through providing users with helpful information with respect to their current location. As a recent trend, virtual location-based services consider webpages or sites associated as 'virtual locations' that online users can visit. The presence of links between virtual locations and the corresponding physical locations (e.g., geo-location information of a restaurant linked to its website), allows for novel types of services and applications which constitute virtual location-based services (VLBS). Their success largely depends on the existence of websites referring to physical locations. In this paper, we investigate the usefulness of linking virtual and physical locations. For this, we analyze the presence and distribution of virtual locations, i.e., websites referring to places, for two Irish cities. Using simulated tracks based on a user movement model, we investigate how mobile users move through the Web as virtual space. Our results show that virtual locations are omnipresent in urban areas, and that the situation that a user is close to even several such locations at any time is rather the normal case instead of the exception.**

***Keywords*-virtual location-based services, case study, user movements, evaluation**

I. Introduction

With the advances in mobile technologies, people can access the Web almost everywhere. Modern mobile devices are able to determine the physical context of a user such as the current location. These developments have spurred the success and popularity mobile applications which provide mobile users with information based on their physical context. The location of users is one of the most important physical context as evident by the huge success of location-based services such as Foursquare, Google Places etc.

In previous works [19], [21], we motivated and studied the advantages of linking locations with the corresponding websites towards the notion of virtual location-based services (VLBS). VLBS consider a webpage, sets of pages or complete websites representing a location as a *virtual location*. For example, given a location like the computer science building of a university, the website for that computer science department denotes a virtual location. VLBS utilize this idea to provide additional information about (virtual) locations or enable users to communicate and collaborate between others being physically on-site. In [21], we merge both types of location-based services by connecting physical with virtual locations in an overarching model of "space".

In this paper, we investigate the feasibility, usefulness and potential benefits of VLBS in detail. The success of VLBS depends on the widespread existence of links between physical and virtual locations. More specifically, we formulate the following research questions: (a) How common are locations with a physical as well as a virtual representation and how are they distributed across cities? (b) How do mobile users in the real world move through the virtual space in terms of being close to virtual locations? (c) What are new insights into the realization of VLBS to improve the online experience of both web and mobile users?

For answers, we provide an in-depth case study analyzing the existence and distribution of virtual locations. We focus on locations which "naturally" feature a physical counterpart such as websites dedicated to physical locations, e.g., the websites of shops, hotels, bars, etc. VLBS enable novel functionalities to users such as context-aware interactions or computer-supported cooperative work on the Web. For example, a Web user at home and browsing a restaurant's website might contact mobile users present in the restaurant to inquire about the current occupancy (geo-social search). A mobile user walking by a shop can be notified about the shop's website with its current offers (mobile web advertising). Also new ideas for mobile gaming involving both Web and mobile users are conceivable.

For our case study, we collected virtual locations within the cities of Dublin and Galway, Ireland. We use the real-world user visits on Foursquare (https://foursquare.com/) to simulate user movements in both cities to investigate mobile users' visits at virtual location while moving through virtual space. We distinguish between two types of movements: *recurring movements* refer to users' movements that are part of their daily routine such as going to work; *non-recurring movements* refer to less common movements, e.g., going to a pub after watching a movie in a theatre. Our results show that in urban environments virtual locations are virtually omnipresent. Furthermore, our simulation of user movements show that most of the physical user movements within a city result in users traversing through many virtual locations. This is particularly true for non-recurring movements, since here mobile users are more likely to pass areas with a high density of virtual locations. These results promote the success of real-world VLBS applications.

Paper outline: Section II reviews related work. Section III provides a basic model of the virtual space and presents our VLBS prototype. Section IV describes the collection and generation of the dataset we used within our evaluation. Section V presents the results of our comprehensive case study. Finally, Section VI concludes and birefly outlines on-going work and challanges.

II. Related Work

Location-based services. Location-based services gained enormous popularity since mobile devices enabled users to get contextualized information based on their location. The locations of users are sensitive information which users are typically not willing to share. Existing approaches to preserve users' privacy aim to not disclose one's exact location but rather an estimate [6], [16]. The main challenge here is identifying a meaningful trade-off between the level of privacy and the quality of the provided service. Various user studies have been conducted to investigate users' preferences with respect to sharing their location with others, e.g., [2], [17]. Besides privacy, other user studies such as [3], [5] investigate the effects of multiple factors (e.g., costs, security, quality) on the successful adoption of such services.

Location-based social networks. Location-based social networks enable users to establish social connections with others and express their visits to places along with their social profiles. Functionalities such as checking-in at places, rating or commenting on them and commenting are very user-centric as they also bring social contexts into consideration. Popular social networks are Facebook Places, Foursquare, and Google Places. Existing works analyzed the user visits to places and the effect of social ties between users on the user movement patterns [4], [14], [24]. The results in [1] indicate that social ties of users can be used to discover approximate locations of users. [15], [22] show that user mobility patterns can be used to predict the social ties between users. The results show strong links between the social network of a user and his/her movements.

Towards virtual location-based services. The concept of virtual locations originates from the efforts towards collaboratively browsing and searching the Web. SearchTogether [8] and CoScripter [7] enable collaborative browsing between users working with their own computers. PlayByPlay [23] demonstrates the use of collaborative browsing with a system which lets the mobile device users and Web users collaborate and communicate. COBS (*CO*llaborative *B*rowsing and *S*earching) [19], [20] proposes a browser extension providing a proof-of-concept implementation that allows users visiting the same site to communicate with each other. In [21], we present a novel approach to enable the communication between users visiting virtual locations and users present at physical locations by linking virtual locations, i.e., websites, to physical locations.

In summary, traditional location-based services and related services on the Web have been investigated independently. In [21], we have proposed a framework to link the physical locations to their virtual locations in order to enable better communication and collaboration between users. We have also presented some preliminary results regarding its potential benefits. In this paper, analyzing openly available datasets, we show the usefulness for merging physical and virtual locations in order to develop novel VLBS.

III. Virtual Location-Based Services

We introduce the notion of VLBS as follows: We first provide a model for the virtual space, and then present VLIMSy, our current VLBS prototype.

A. A Model of the Virtual Space

Our approach is to adopt the notion of a user's location from the real world to the Web. In the following, we define the required concepts of a virtual coordinate and virtual location. We limit the presentation of the model of the virtual space to the concepts required for this paper; we present the full model in [21]. Simply speaking, space describes the possibilities where a person "can be". Given these notions, we define the virtual space as the set of web pages a user can visit. In geographic terms, the most fine-grained way to specify a mobile user's current position is by means geo coordinates, e.g., longitude and latitude. Mapping the concept to the virtual space, the current position of a user is the web page the user is visiting. Thus, within our framework, each page is uniquely identified by a URL.

Definition 1 (*Virtual coordinate*). A virtual coordinate *vc* is the URL of a webpage. ■

Typically, not the distinct page but the category or topic or similar concepts of a page are of interest to describe a web user's location. We therefore extend the definition of a virtual location beyond a single virtual coordinate.

Definition 2 (*Virtual location*). A virtual location *vl* is a distinct, non-empty, finite set of virtual coordinates $\mathcal{V} = \{vc_1, vc_2, ..., vc_n\}$, with $\mathcal{V}_1 \cap \mathcal{V}_2 = \emptyset$. ■

The set of virtual coordinates that constitutes a virtual location is application-specific. Throughout this paper, we use the domain of a URL as identifier of a virtual location, i.e., we group all subpages of a website into one location. This is a reasonable assumption for websites associated to physical locations such as hotels, shops, businesses, etc., which are in the scope of our evaluation.

B. A Simple VLBS Application

Our proof-of-concept implementation of a VLBS, called VLIMSy, allows the exchange of presence information and messages of web and mobile users based on their physical and/or virtual location. For example, a web user browsing a shop's website can connect with other web user visiting the same site or mobile users close to the shop.

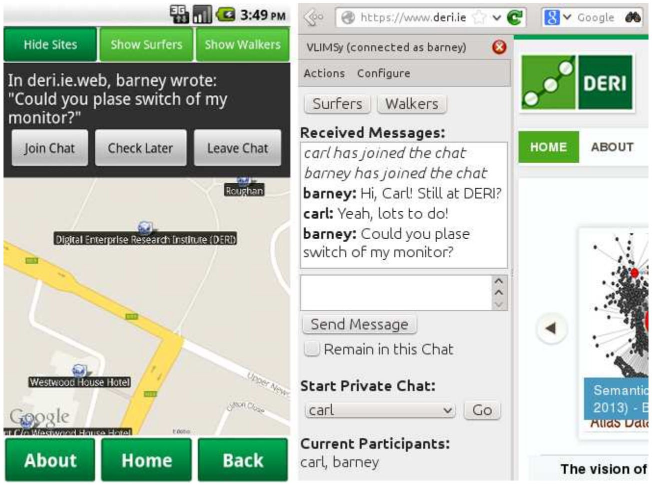


Figure 1. Screenshots of the mobile application and browser add-on.

Backend Architecture. A data repository maintains the mapping between the physical and virtual locations. We focus on "single point" locations like hotels, pubs, shops, etc. We store areas like parks or golf courses also using single geo coordinates. We provide the presence and instant messaging service based on the open-standard eXtensible Messaging and Presence Protocol (XMPP, http://xmpp.org). For VLIMSy, the most relevant concept is the "group chat"; we assign each location to a group chat. The intuition is that users at the same location are in the same group chat and are therefore aware of each other. In general, the physical and virtual locations of a user differ. For example, a customer in a shop is not necessarily browsing the shop's website. We therefore distinguish between a *geo group chat* representing the physical representation and a *web group chat* representing virtual representation of a location. Besides presence information, we also make use of the possibility to exchange messages. We support group chats and the user-to-user communication between web and mobile users.

Frontend Applications. Given the different devices and applications for web users (e.g., at home) and mobile users, we provide two interfaces to interact with VLIMSy: a web browser add-on and a mobile application; see Figure 1.

Web browser add-on. We aim for a seamless integration of our presence mechanisms into the normal browsing experience of users. We therefore implemented a browser add-on featuring a sidebar to provide our instant messaging service. The add-on maintains an XMPP connection with references to two group chats, (a) to the web group chat of the currently visited website, and (b) to the geo group chat of the corresponding physical location (if available). The latter enables web users to be aware of mobile users that are close to the physical location of the visited website. Using the sidebar, user can engage in private and group chats with other users.

Mobile phone application. We implemented an Android application with two main features: A map based on the GOOGLE MAPS API displays all available virtual locations, web and mobile users in the vicinity as different markers. Clicking on a marker displays some basic information about the corresponding location, web or mobile user. This information window also allows a mobile user (a) to enter the group chat of a virtual location or (b) send a private message to other web or mobile users. The second feature is a basic chat client for private and group chats. We assume that a mobile device is capable of determining its location. If a mobile users is close to virtual location, the application automatically enters the corresponding geo group chat, and thus making the mobile user visible to web users.

VLIMSy is our current experimental setup to illustrate the potential of VLBS and to get deeper insights into the challenges of their implementation as real-world applications. In principle, our setup can easily be extended to meet the needs of users such as setting their preferences in terms of privacy, access control, visibility radius, etc. Further meaningful extensions comprise mechanisms towards trust and reputation management to incentivize users to participate as well as to discourage malicious behavior.

IV. DATA COLLECTION AND GENERATION

This section describes the collection and generation process of the data we used for our evaluation.

A. Physical and Virtual Locations Data

Our dataset consits of the physical and virtual locations for two cities in Ireland, Dublin and Galway. We used the `Google Places API`[1] to find all places within the city limits of Dublin and Galway. For only evaluation, we considered only places that feature a website, i.e., a virtual location. As a result, we collected approximately 1,400 entries for Galway and 16,400 entries for Dublin with each entry featuring both a physical and virtual location. We made all data used in our evaluation available on an accompanying website.[2]

B. Simulated User Movements

We simulated user movements across Dublin and Galway. User movement has been analyzed using the user check-in activities on location based social networks [9], [10]. These works have shown that movements of any user occur within a specific geographical area with occasional movements outside the area. It has been demonstrated that many of the user movements have repeatability such as travelling to work place and users rarely travel between any random locations [11], [12]. The model advocating this is known as *Activity-Based Travel Demand Modeling* and has been extensively used to model users' travel decisions.

[1]https://developers.google.com/places/
[2]http://vmusm02.deri.ie/vlimsy

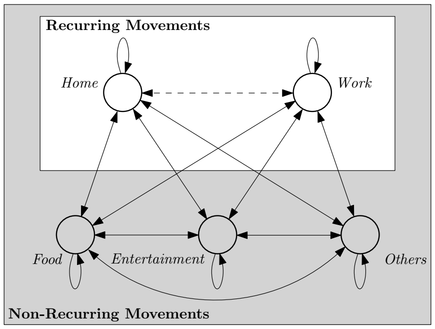


Figure 2. User movements between various categories of places

Algorithm 2 Simulation of non-recurring movements

1: select a place category c_{start} as $c_{start} \sim Uniform(1/|C_{all}|)$
2: **if** c_{start} is home **then**
3: select l_{start} as $l_{start} \sim Uniform(1/|H|)$
4: **else**
5: $\pi_{start} \sim Dirichlet(\alpha_{c_{start}})$
6: select l_{start} as $l_{start} \sim Discrete(\pi_{start})$
7: **end if**
8: select a place category c_{end} as $c_{end} \sim Discrete(w_{c_{start}})$
9: **if** c_{end} is home **then**
10: select l_{end} as $l_{end} \sim Uniform(1/|H|)$
11: **else**
12: $\pi_{end} \sim Dirichlet(\alpha_{c_{end}})$
13: select l_{end} as $l_{end} \sim Discrete(\pi_{end})$
14: **end if**
15: return (l_{start}, l_{end})

To simulate users' start and end locations, we used the total number of user check-ins a place has on Foursquare. Places in Foursquare belong to categories such as 'restaurant', 'office', etc. We classified the place categories as *Home*, *Work*, *Food*, *Entertainment* and *Others* (cf. [18]). A path comprises of a user moving between places of different categories. We distinguish two types of user movements: *recurring movements* between *Home* and *Work* represent the weekday routine of users; all other movements we denote as *non-recurring movements*. Figure 2 illustrates this, with *Home* and *Work* and the dashed edge reflecting recurring movements; non-recurring movements consider all five categories but exclude paths between *Home* and *Work*.

Algorithm 1 Simulation of recurring movements

1: select a home location l_{start} as $l_{start} \sim Uniform(1/|H|)$
2: select a work place l_{end} as follows
3: $\pi_W \sim Dirichlet(\alpha_W)$
4: select l_{end} as $l_{end} \sim Discrete(\pi_W)$
5: return (l_{start}, l_{end}) and (l_{end}, l_{start})

Algorithm 1 describes the way samples of start and end locations are generated to simulate the recurring user movements. H and W denote the set of all locations belonging to *Home* and *Work* respectively. A uniform sampling over H with the probability of $1/|H|$ ensures that every home location has equal probability of being start location l_{start} (Line 1). We select the location l_{end} belonging to *Work* based on the total number of check-ins. We use the Dirichlet prior for a discrete distribution motivated by Bayesian Bootstrapping [13] to smooth the distributions as we have small datasets in terms of number of places per category and number of check-ins, compared to the number real-world locations and user visits. We first sample a distribution function π_W from a Dirichlet distribution with the parameters $\alpha_W = (\alpha_1, \ldots, \alpha_w)$, where $w = |W|$ and α_i is the total number of user check-ins at the i^{th} work place (Line 3). This ensures that the discrete distribution sampled from $Dirichlet(\alpha_W)$ favors the selection of a work place as end location with a higher number of check-ins. Finally, we sample l_{end} from a discrete distribution with the sample space W and distribution function π_W (Line 4).

Algorithm 2 describes the way samples of l_{start} and l_{end} are chosen to simulate the non-recurring user movements. With C_{all} being the set of all five categories of places, a uniform sampling is carried out over C_{all} to obtain the category c_{start} of any start location (Line 1), so that all categories are chosen equally. If c_{start} is *Home*, we select any home location with equal probability (Line 3) to make sure that all home locations are well-represented in the simulations. Otherwise, we again use the Dirichlet prior for the discrete distribution to select l_{start} to reflect the number of check-ins (Lines 5-6). The selection of l_{end} comprises two steps: Firstly, we select a category c_{end} based on the choice of c_{start}, favoring categories with many locations. And secondly, we select l_{end} as location of category c_{end}, favoring locations with many check-ins. To select c_{end}, we use a stochastic transition matrix defined as:

$$\begin{array}{l} \\ Home \\ Work \\ Food \\ Entertainment \\ Others \end{array} \begin{array}{c} \begin{array}{ccccc} Home & Work & Food & Entertainment & Others \end{array} \\ \begin{pmatrix} \varepsilon & 0 & w_{13} & w_{14} & w_{15} \\ 0 & \varepsilon & w_{23} & w_{24} & w_{25} \\ w_{31} & w_{32} & \varepsilon & w_{34} & w_{35} \\ w_{41} & w_{42} & w_{43} & \varepsilon & w_{45} \\ w_{51} & w_{52} & w_{53} & w_{54} & \varepsilon \end{pmatrix} \end{array}$$

For non-recurring movements, the transition probabilities between *Home* and *Work* are 0. We calculate the transition probabilities between categories as $w_{ij} = \frac{N_j}{M_i} - \frac{\varepsilon}{Z_i}$ with $i \neq j$. N_j is the number of distinct places belonging to j^{th} category of the matrix. $M_i = \sum_k N_k$ where k is any column index whose entry is not assigned with ε or 0 in the i^{th} row. Z_i is the number of categories to which a transition from the i^{th} category can be made. Hence, $Z_1 = Z_2 = 3$ and $Z_3 = Z_4 = Z_5 = 4$. ε denotes the self-transition probability of start and end location belonging to the same category. Setting the ε to a small value ensures that movements such as going from one *Food* place another are rare. Setting the w_{ij} values based on the number of places belonging to various categories ensures that the most visited categories are favored in the simulation. Given c_{start}, we choose the end location category c_{end} from a discrete distribution with the distribution function defined by the row vector $w_{c_{start}}$. Once we have selected c_{end}, we sample l_{end} in a fashion similar to the selection of the start location (Lines 9-14).

r_v	*radius of vicinity* of virtual locations: minimum distance (in meter) between users and locations to be considered as visits of the users at locations.
t_v^{min}	*minimum visiting time*: minimum time (in seconds) a user has to spend in the vicinity radius of a virtual location to be considered as visit of users at locations.
l_{max}	*maximum path length*: maximum length in kilometer of simulated paths.

Table I
LIST OF EVALUATION PARAMETERS

	Home	*Work*	*Food*	*Entertainment*	*Others*
Dublin	4413	1280	1425	634	2269
Galway	417	228	275	156	442

Table II
NUMBER OF PLACES BELONGING TO DIFFERENT CATEGORIES

	Home	*Work*	*Food*	*Entertainment*	*Others*
Dublin	0.184	0.158	0.232	0.322	0.104
Galway	0.106	0.181	0.241	0.323	0.149

Table III
STATIONARY DISTRIBUTIONS OF CHECK-INS PER CATEGORY

Finally, for each selected start and end location, we used the `Google Directions API`[3] to obtain the paths between the two locations. Since we focus on walking users , we used `walking` as travel mode to get the directions via pedestrian paths and side-walks (where available). Note that with this method we consider direct paths between locations.

V. EVALUATION

In this section, we present the results of our case study, i.e., whether merging the physical and virtual space involves a sufficient overlap to be of practical relevance. Table I describes the parameters we considered within the analysis.

A. Simulated User Movements

We crawled places and check-in activities for Dublin and Galway on Foursquare (see Table II) to obtain the parameters of the various probability distributions in our simulation model. Table III shows the stationary distributions for the five different categories, obtained by computing the long-term behavior of the Markov chain defined by the transition matrix defined in Section IV-B. We set $\varepsilon = 0.1$ as the self-transition probability parameter for the transition matrix. We found that the stationary distributions of user check-ins are very similar for Dublin and Galway and also are similar to the result reported in the previous studies [15], [18].

To obtain a reasonable number of user movements, we first generated 5,000 paths for recurring and non-recurring user movements. Figure 3 shows the distribution of path

[3]https://developers.google.com/maps/documentation/directions/

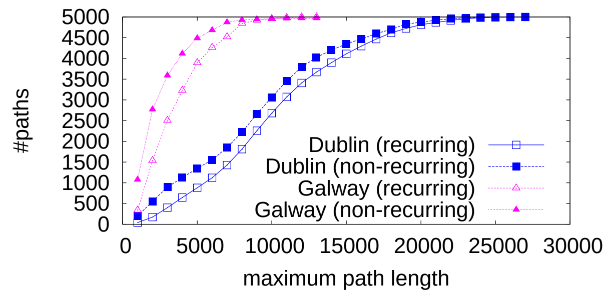


Figure 3. Distribution of path lengths of the simulated user movements

	Dublin	**Galway**
population (in 2011)	525,400	75,500
#places	39,237	3,692
#virtual locations	16,485 (42.0%)	1455 (39.4%)

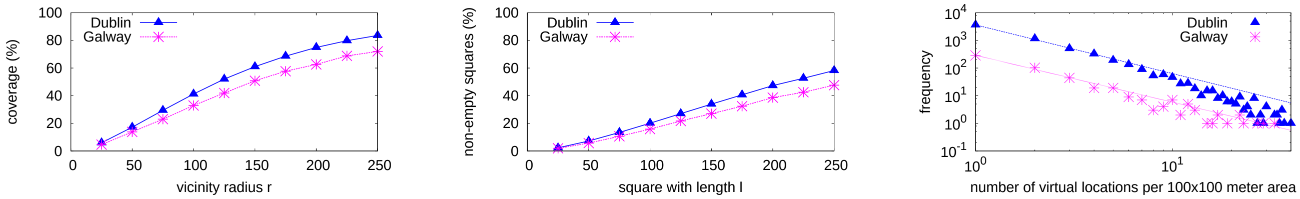
Table IV
BASIC STATISTICS OF COLLECTED DATA

lengths, i.e., the number of paths with lengths $\leq l_{max}$. As expected, given its much larger size, paths in Dublin are on average much longer than in Galway. Furthermore, popular locations for non-recurring movements are much more concentrated within the city of Galway. Also, paths for recurring movements are on average longer than for non-recurring movements, since *Home* places are typically in areas with less work virtual locations. Considering walking users, overlong paths are not meaningful. Hence, for experiments with a fixed maximum path length we set $l_{max} = 3$km. Otherwise, we vary l_{max} between 1 and 5 kilometers.

B. Coverage & Distribution of Locations

Table IV shows the population size, number of places and virtual locations (i.e., places that feature a website) we collected for Dublin and Galway. While both cities differ regarding their population size, the number of places is roughly proportional to the size. Moreover, the number of places that feature a website is for both cities about 40%.

We first calculated the coverage in percent; see Figure 4(a). Naturally, the coverage increases for larger r_v, resulting in up to 83% (72%) coverage for Dublin (Galway) for $r_v = 250$m. Regarding the distribution of virtual locations, we divided the areas of the cities into squares with different side lengths l and counted the number of virtual locations within each square. Figure 4(b) shows the ratio of non-empty squares, which naturally increases for larger squares. Empty squares typically cover city parks or purely residential areas. Figure 4(c) shows the distribution all non-empty squares for $l = 100$m. Not unexpectedly, the number of virtual locations per square and their respective frequency show a power-law relationship: While most squares contain only a small set of locations, few squares contain a very large number of locations (e.g., city centers, business parks).

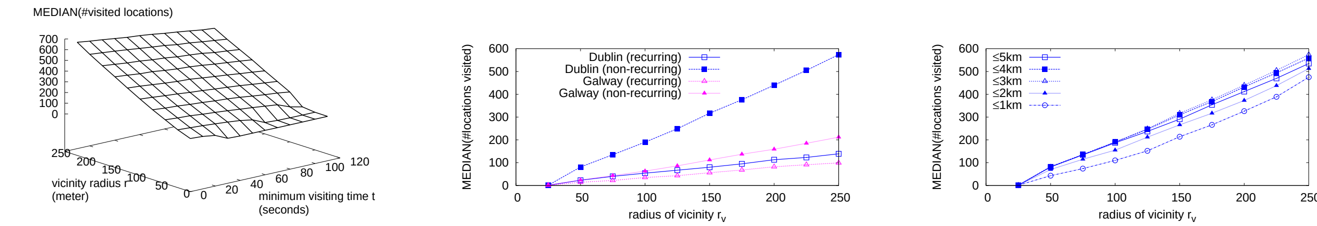


(a) coverage for different vicinity radiuses r

(b) ratio of non-empty squares

(c) distribution of non-empty squares

Figure 4. Coverage and distribution analysis regarding virtual locations across the cities of Dublin and Galway

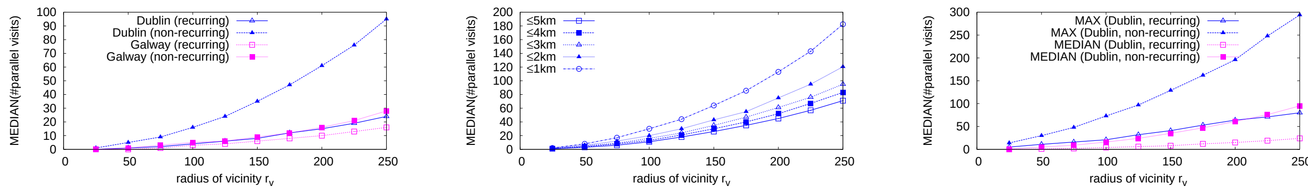


(a) Results for different values of r_v and t_v^{min} (Dublin/non-recurring, $l_{max} = 3km$)

(b) Results for each path datasets of Dublin and Galway ($t_v^{min} = 60$s, $l_{max} = 3km$)

(c) Results for different maximum path lengths l_{max} ($t_v^{min} = 60$s, Dublin/non-recurring)

Figure 5. Average (median) number of visited locations



(a) Results for different different graph dataset ($l_{max} = 3km$)

(b) Results for different maximum path lengths l_{max} ($t_v^{min} = 60$s, Dublin/non-recurring)

(c) "Median of medians" vs. "median of maximums" ($l_{max} = 3km$, Dublin)

Figure 6. Average number of parallel visits

C. Overlap of Physical and Virtual Space

Average number of visited locations. With number of visited locations being rather skewed, we use the median to quantify the average number of visited locations. Figure 5(a) shows the results for the non-recurring movements in Dublin for various r_v and t_v^{min} and for $l_{max} = 3$km. The results for the other path datasets differ only in the absolute values. Most naturally, the smaller r_v and the larger t_v^{min} the less locations a user is visiting. The vicinity radius has the greater effect on the results due to the low speed (average walking speed). Overall, the results indicate that visiting many virtual locations while walking through a city is very common.

Next, we looked at the difference between the path datasets (Figure 5(b)). As expected, the paths derived from non-recurring movements in Dublin cross the most virtual locations, since they typically pass through areas with many locations. Recurring paths, on the other hand, often originate outside such areas. For Galway, the differences between recurring and non-recurring paths are also visible. However, the results are much lower than for Dublin since the areas with many virtual locations are smaller. We therefore argue that VLBS are more likely to be successful in larger cities.

We also varied l_{max} to see its effect. As Figure 5(c) shows, the effect of varying l_{max} is limited. The average number of visited locations even slightly drops again for increasing l_{max}. Our explanation is that longer paths are more likely to cover areas with a lower location density. In this sense, the maximum path length resulting in the highest average number of visited locations vaguely indicates the size of high-density areas (in case of direct paths).

Average number of parallel visits. We performed several experiments to measure the average number of parallel visits. For each individual path the number of parallel visits changes over time, and calculated the median to represent the average number of parallel visits for each path. We eventually calculated the median of medians to report the average number of visits across our path datasets.

Figure 6(a) shows the results for our path datasets. Not surprisingly, the number of parallel visits correlate with the overall number of visited locations. Even for moderate r_v, a mobile user being close to multiple virtual locations at the same time is very common. We then compared the results regarding the number of parallel visits for different l_{max}. Figure 6(b) shows for non-recurring movements in Dublin that for small l_{max} the number of parallel visits increases. This is, again, due to the high probability of short paths crossing areas with many virtual locations.

Finally, Figure 6(c) compares the median of medians results with the ones for median of maximums for the two Dublin path datasets. For the median of maximums, we calculated the maximum number of parallel visited locations (instead of the median) for each path. Naturally, the median of maximums is significantly higher than the median of medians, resulting in up to several hundred parallel visits for non-recurring paths in Dublin. From an application perspective this means that Web and mobile user potentially find themselves in the presence of many other users visiting the same locations at the same time. While this is, in general, a worthwhile situation, it also poses new challenges, e.g., to avoid unmanageable number of parallel chats in VLIMSy.

Accumulated visiting time. In our last series of experiments, we evaluated the overlap between the physical and virtual spaces in terms of the average accumulated visiting time. This value represents the time a mobile user was close to virtual locations independent of parallel visits.

We first calculated the accumulated times for different r_v and t_v^{min}. Figure 7(a) shows the results for the non-recurring movements in Dublin and $l_{max} = 3$km, using the median to average over the path datasets. Again, the effect of t_v^{min} is less pronounced than the one of r_v due to the slow travel speed. Most striking, however, is the long accumulated visiting times of several hours. The reason for that is the often high number of virtual locations a mobiles user is close-by at the same time. Figure 7(b) shows the results for all path datasets and $l_{max} = 3$km. As expected, the results are lower for paths in Galway than in Dublin. However, even for recurring movements in Galway, the average accumulated visiting time goes up to several hours. Thus, although mobile users walk only for a short time – not more than 36 minutes given a walking speed of 5km/h and $l_{max} = 3$km – they are typically significantly longer present in the virtual space.

Finally, we investigated the effect of the maximum path length l_{max} on the average accumulated visiting time; see Figure 7(c). Again, longer maximum paths length do not result in longer visiting times since long paths tend to (partly) cross areas with less virtual locations. This is a result of our experiment setup using direct paths with a constant travel speed, and hence might differ for different settings. In general, our setup yields the "worst" results in terms of the accumulated visiting time, since we, for example, do not consider any breaks, e.g., for shopping or eating.

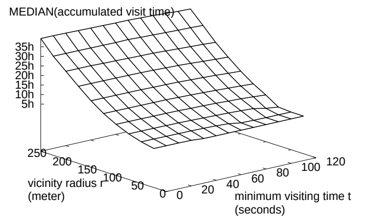
D. Discussion of Results

Potential benefits of VLBS. To be useful, VLBS require that the physical and virtual space overlap sufficiently. Our results consistently confirm this for urban areas. Areas like city centers feature such a high density of virtual locations that users are close to a large number of locations at any time. Thus, we deem the virtual presence of mobile users of practical importance for the design and development of new kinds of VLBS bridging the physical and virtual world. Although our results are already very promising, we expect a more significant overlap in real-world deployments. So far, we considered only the case where mobile users walk on straight paths. Imagine, however, a tourist strolling through a city center: His/her path might cross itself multiple times, the walking speed is slow and includes breaks for shopping or eating. This further increases the overlap between the virtual and physical world, particularly regarding the time spent close-by virtual locations.

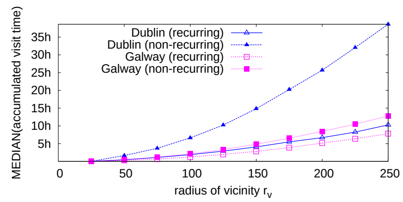
Impact on implementations of (virtual) location-based services. In areas with a high density of virtual locations, such as city centers, the probability that a mobile user is present at a large number of virtual locations at the same time is very high. Depending on the number of already present users this can lead to an unmanageable number of parallel encounters, and therefore might negatively affect the user experience. Thus, suitable mechanisms limiting a user's presence to a reasonable number of parallel virtual locations are required e.g., filtering or ranking techniques. Meaningful techniques consider the different absolute distances or more sophisticated parameters (e.g., how often a user is close a particular location) to determine the ranking or filter conditions. A further approach is to dynamically adapt important system parameters such as the vicinity radius r_v or the minimum visiting time t_v^{min}.

VI. Conclusions

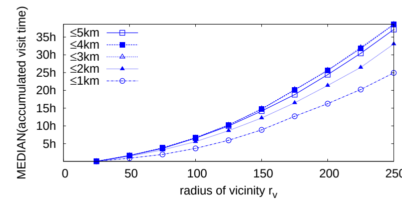
Merging virtual locations (i.e., web pages or sites) with points of interest in the real world, opens new opportunities for the design and development of novel virtual location-based services. Mobile users can get additional information about websites associated with nearby locations provided on their devices. Users at home browsing a website can get in contact with people on-site, i.e., users that are close to the location connected to the visited website. In this paper, we showed that such approaches promise to have real practical impact.Firstly, virtual locations are virtually omnipresent covering very large portions of urban areas, with the distribution of virtual locations showing power-law relationships. And secondly, using simulated tracks describing different categories of user movements, we demonstrated that the situation where users are nearby such locations is rather the normal than the exceptional case. Moreover, as a result of the distribution of virtual locations, users are often close to many of them at the same time.



(a) Results for different values of r_v and t_v^{min} (Dublin/non-recurring, $l_{max} = 3km$)



(b) Results for different different graph dataset ($l_{max} = 3km$)



(c) Results for different maximum path lengths l_{max} ($t_v^{min} = 60s$, Dublin/non-recurring)

Figure 7. Accumulated visiting time

Summing up, our results show that there is a significant overlap between the physical and virtual space promoting the practical relevance and potential benefits of VLBS. Furthermore, the results also serve as input for ideas and the design and implementation of such services. Towards the design and implementation real-world VLBS we see various immediate challanges. Most importantly, like for traditional location-based services, privacy is a relevant issue since the website a user is browsing on may represent sensitive information. Further challanges are, depending on the specific application, incentivizing users to share the (virtual) locations and to provide user-generated content. We argue that not only existing techniques from traditional location-based services – e.g. privacy preservation techniques obfuscating users' exact location – but also mechanisms from other related fields, such as online social networks as well as reputation and recommender systems, are worth investigating.

References

[1] L. Backstrom, E. Sun, and C. Marlow. Find me if you can: improving geographical prediction with social and spatial proximity. In *WWW '10*. ACM, 2010.

[2] A. B. Brush and Krumm. Exploring End User Preferences for Location Obfuscation, Location-Based services, and the Value of Location. In *Ubicomp '10*. ACM, 2010.

[3] S. Chang, Y.-J. Hsieh, T.-R. Lee, C.-K. Liao, and S.-T. Wang. A User Study on the Adoption of Location Based Services. In *Advances in Web and Network Technologies, and Information Management*. Springer, 2007.

[4] E. Cho, S. A. Myers, and J. Leskovec. Friendship and mobility: user movement in location-based social networks. In *SIGKDD '11*. ACM, 2011.

[5] I. Junglas. An Experimental Investigation of Location-Based Services. In *HICSS '05*, 2005.

[6] J. Krumm. A Survey of Computational Location Privacy. *Personal and Ubiquitous Computing*, 13(6):391–399, 2009.

[7] G. Leshed, E. M. Haber, T. Matthews, and T. Lau. CoScripter: Automating & Sharing How-to Knowledge in the Enterprise. In *CHI '08*. ACM, 2008.

[8] M. R. Morris and E. Horvitz. SearchTogether: An Interface for Collaborative Web Search. In *UIST '07*, 2007.

[9] T. Nguyen and B. K. Szymanski. Using location-based social networks to validate human mobility and relationships models. In *ASONAM '12*. IEEE, 2012.

[10] A. Noulas, S. Scellato, R. Lambiotte, M. Pontil, and C. Mascolo. A tale of many cities: universal patterns in human urban mobility. *PloS one*, 7(5):e37027, 2012.

[11] N. Oppenheim et al. *Urban travel demand modeling: from individual choices to general equilibrium.* John Wiley and Sons, 1995.

[12] E. I. Pas. Recent advances in activity-based travel demand modeling. In *Activity-Based Travel Forecasting Conference*, 1997.

[13] D. B. Rubin. The bayesian bootstrap. *The annals of statistics*, 9(1), 1981.

[14] S. Scellato, A. Noulas, R. Lambiotte, and C. Mascolo. Socio-spatial properties of online location-based social networks. *ICWSM*, 11:329–336, 2011.

[15] S. Scellato, A. Noulas, and C. Mascolo. Exploiting place features in link prediction on location-based social networks. In *SIGKDD '11*. ACM, 2011.

[16] K. Shin, X. Ju, Z. Chen, and X. Hu. Privacy Protection for Users of Location-Based Services. *Wireless Communications, IEEE*, 19(1):30–39, 2012.

[17] E. Toch, J. Cranshaw, P. H. Drielsma, J. Y. Tsai, P. G. Kelley, J. Springfield, L. Cranor, J. Hong, and N. Sadeh. Empirical Models of Privacy in Location Sharing. In *Ubicomp '10*. ACM, 2010.

[18] M. A. Vasconcelos, S. Ricci, J. Almeida, F. Benevenuto, and V. Almeida. Tips, dones and todos: uncovering user profiles in foursquare. In *WSDM '12*. ACM, 2012.

[19] C. von der Weth and A. Datta. COBS: Realizing Decentralized Infrastructure for Collaborative Browsing and Search. In *AINA '11*. IEEE, 2011.

[20] C. von der Weth and A. Datta. FAST: Friends Augmented Search Techniques - System Design {&} Data-Management Issues. In *WI-IAT '11*. IEEE, 2011.

[21] C. von der Weth and M. Hauswirth. Finding Information Through Integrated Ad-Hoc Socializing in the Virtual and Physical World. In *WI-IAT '13*. IEEE, 2013.

[22] D. Wang, D. Pedreschi, C. Song, F. Giannotti, and A.-L. Barabasi. Human mobility, social ties, and link prediction. In *SIGKDD '11*. ACM, 2011.

[23] H. Wiltse and J. Nichols. PlayByPlay: Collaborative Web Browsing for Desktop and Mobile Devices. In *CHI '09*, New York, NY, USA, 2009. ACM.

[24] Y. Zheng, X. Xie, and W.-Y. Ma. Geolife: A collaborative social networking service among user, location and trajectory. *IEEE Data Eng. Bull.*, 33(2):32–39, 2010.