

Exploiting Behavioral User Models for Point of Interest Recommendation in Smart Museums

Seyyed Hadi Hashemi and Jaap Kamps

University of Amsterdam, Amsterdam, The Netherlands,
Email: {hashemi,kamps}@uva.nl

ABSTRACT

The Internet of Things (IoT) holds the promise to blend real-world and online behavior in principled ways, yet we are only beginning to understand how to effectively exploit insights from the online realm into effective applications in smart environments. Such smart environments aim to provide an improved, personalized experience based on the trail of user interactions with smart devices, but how does recommendation in smart environments differ from the usual online recommender systems? And can we exploit similarities to truly blend behavior in both realms to address the fundamental cold-start problem? In this article, we experiment with behavioral user models based on interactions with smart devices in a museum, and investigate the personalized recommendation of what to see after visiting an initial set of Point of Interests (POIs), a key problem in personalizing museum visits or tour guides, and focus on a critical one-shot POI recommendation task—where to go next? We have logged users’ onsite physical information interactions during visits in an IoT-augmented museum exhibition at scale. Furthermore, we have collected an even larger set of search logs of the online museum collection. Users in both sets are unconnected, for privacy reasons we do not have shared IDs. We study the similarities between users’ online digital and onsite physical information interaction behaviors, and build new behavioral user models based on the information interaction behaviors in i) the physical exhibition space, ii) the online collection, or iii) both. Specifically, we propose a deep neural multi-layer perceptron (MLP) based on explicitly given users’ contextual information, and set-based extracted features using users’ physical information interaction behaviors and similar users’ digital information interaction behaviors. Our experimental results indicate that the proposed behavioral user modeling approach, using both physical and online user information interaction behaviors, improves the onsite POI recommendation baselines’ performances on all evaluation metrics. Our proposed MLP approach achieves 83% precision at rank 1 on the critical one-shot POI recommendation problem, realizing the high accuracy needed for fruitful deployment in practical situations. Furthermore, the MLP model is less sensitive to amount of real world interactions in terms of the seen POIs set-size, by backing off to the online data, hence helps address the cold start problem in recommendation. Our general conclusion is that it is possible to fruitfully combine information interactions in the online and physical world for effective recommendation in smart environments.

KEYWORDS

Human information interaction, Onsite logs, Behavioral user models, POI recommendation, Internet of things

1. Introduction

The last decade witnessed a surge of interest in the implementation of Internet of Things (IoT) in different applications, such as smart shopping malls and smart mu-



Figure 1. Interactive POIs in a museum physical space, consisting of a series of pedestals with screens and actuators integrated into the Roman department of the Allard Pierson Museum of Archaeology in Amsterdam, The Netherlands.

seums, which provide the infrastructure for understanding users' physical interaction behavior and consequently their preferences in interacting with smart environments (Atzori, Iera, & Morabito, 2010; Barnaghi, Wang, Henson, & Taylor, 2012; Friess, 2013; Hashemi, Hupperetz, Kamps, & van der Vaart, 2016; Hashemi & Kamps, 2017a; Hernández-Muñoz et al., 2011; Perera, Zaslavsky, Christen, & Georgakopoulos, 2014). This prompts a range of questions: In what ways can tracking people in their real-life behavior and trying to understanding their interaction behaviors be helpful? Is it possible to give effective recommendations to users by tracking them using IoT but without getting any explicit information, like ratings, about their preferences?

Imagine you are at a huge museum like the Louvre in Paris and you want to explore the museum. Usually, it is impossible to visit every single object in a large museums like the Louvre in one day. Furthermore, freely roaming through the museum is more desirable in comparison to the traditional fixed walking route designed in a non-personalized way. Providing personalized experiences for users is highly valuable in this context and will help them to visit all the interesting objects of the museum according to the user's preferences. In this case, how amazing would it be if a contextual recommender system can tell you accurately what to visit without relying on extensive history or explicit feedback from you?

The emergence of applications like the above leads to interest in logging users' onsite physical information interactions, creating a new and potentially exponentially growing data about physical interaction that resembles current online search engine interaction logs. Although understanding users' search behavior and their information needs based on query logs is well-studied (Chuklin, Markov, & Rijke, 2015; Hashemi, Williams, El Kholy, Zitouni, & Crook, 2018a, 2018b; Wang, Zhang, Tang, Zheng, & Zhao, 2016), to the best of our knowledge, there has not yet been any study on how to understand users' behaviors and their information needs based on similarities between users' onsite physical and online digital information interaction behaviors. The main contribution of this paper is to address this research problem by learning a behavioral user model using both onsite physical and online digital user behaviors.

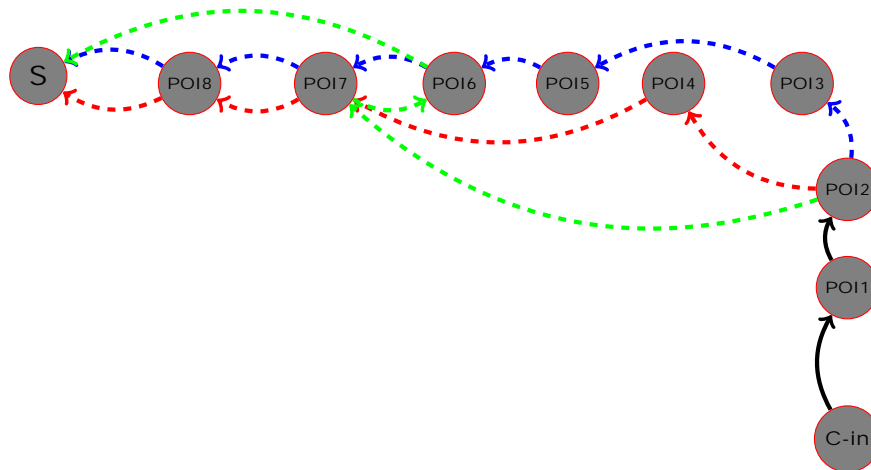


Figure 2. Variance in onsite users' behavior after visiting a set of POIs in a museum exhibition shown in Figure 1. The figure indicates variance of three visitors' preferences in visiting POIs. Each of them shown by a different color, and the black edges are the ones walked by all the three visitors. *C-in* is the check-in station and the *S* is the check-out station.

To this aim, users' onsite physical interactions of visits in a museum and users' online query logs of a search engine on the same collection are logged. Onsite physical information interactions are based on unlocking contents of an installed iPad screen at each POI using RFID tags. For privacy reasons we don't have shared IDs, hence users in both sets are un-connected, and we study the typical cold start case where we have no prior history on a visitor to the smart exhibition in the museum yet we have historical data of users' online interactions with the museum search engine. We study how we can use similarity of users' online and onsite information interaction behaviors with an aim of improving onsite POI recommendation at the smart museum. Figure 1 shows an example of the museum space with the mentioned installations. In this way, we log users' interactions with POIs and track users' visits in the museum. Figure 2 shows the floorplan of an exhibition in a smart museum with an integrated IoT. As it is shown in Figure 2, users behave differently after visiting a set of POIs. The walk-through graph of three real users after checking in at POI₁ and POI₂ is plotted. The blue and red paths show walk-through behaviors of two users tend to check-in at POIs one after the other but with different preferences. The green path shows a user who behaves completely different from the other two and does not check-in at POIs one after the other. This figure shows an example of how different users exhibit different onsite physical behavior, which indicates that understanding and prediction of users' onsite physical behaviors can be challenging and difficult.

Understanding users' onsite physical behavior is also challenging as there are external factors in the environment having impact on users' behavior. As it is studied in (Hashemi, Hupperetz, et al., 2016), users' walk-through behavior and their dwell-time interacting with a POI in an exhibition is affected by the position of the POI in the exhibition. They have also observed a decrease in users interests in interacting with

technology at the end of an exhibition compared to the start of the exhibition. These external factors lead to position and temporal rank bias in the collected onsite sensor logs (Hashemi, Hupperetz, et al., 2016). Furthermore, users' behavior is also affected by other visitors around them, which leads to an observation of crowd-bias in collected onsite interaction logs (Hashemi & Kamps, 2017a).

Such external factors bring an additional complexity to understand users' onsite behavior as it makes users' behavior a combination of "pure" content preferences and other factors like the physical constraints. Moreover, there is a difference in how different users will behave in the presence of external factors as those discussed above. Therefore, understanding users' onsite behavior and preferences in order to provide an effective personalized service in a smart environment is an interesting yet challenging problem. Understanding users' onsite behavior and providing effective personalized POI recommendation become even more challenging in smart museums as in the early stage of launching a smart museum, we do not have access to considerable amount of onsite walk-through sensor logs. Thus, taking advantage of other user preferences signals available for a same collection could be very helpful. To this aim, we study similarity of users' online and onsite preferences by using users' online interaction behavior signals to model their onsite interaction behaviors. Specifically, we build a graph, in which graph nodes are the POIs available in a smart museum and graph edges are created based on users' click-through behavior on an online search engine providing access to the same museum collection. We then define behavioral features based on the built graph, which are used to create our proposed behavioral user models.

In this paper, our main aim is to study the question: *How to model users' information interaction behavior with IoT having an aim of providing a personalized onsite POI recommendation?* Specifically, we answer the following research questions:

- (1) *How to understand users' onsite physical behavior and create a behavioral user model that is able to effectively predict relevant unseen POIs?*
- (2) *How strong are different users' interaction behaviors with IoT in understanding users' preferences?*
 - (a) *Are online digital behaviors similar to onsite physical behaviors? Does understanding online digital users' information interaction behaviors have a positive effect in learning a model to predict unseen relevant POIs and complete users' personalized onsite visits?*
 - (b) *What are the relative importance of each feature extracted based on different users' interaction behaviors in effectiveness of POI recommendation systems?*
- (3) *How effective is behavioral POI recommendation system in one-shot POI recommendation problem?*
- (4) *What is the effect of given seen POIs set-size in the unseen POI recommendation performance?*

This paper builds on and extends the work reported in (Hashemi & Kamps, 2017b) by providing more detail and explanations of the approach and its relation to related work, and further analysis such as a study of the impact of number of seen POIs on the performance of the unseen POI recommendation system. The rest of the paper is organized as follows. In Section 2, we review related work on recommender systems and their use in the museum domain, as well as on tracking behavior in smart environments. Our proposed onsite POI recommendation approach is detailed in Section 3. The experimental setup and results are discussed in Section 4 and 5. In Section 6, we discuss potential future directions of our study in this paper. Finally, we present the

conclusions and future work in Section 7.

2. Related Work

In this section, we discuss related work on context-aware recommendation systems, POI recommendation systems, recommendation systems in museums, and the Internet of Things (IoT).

2.1. Context-Aware Recommendation Systems

Traditionally, recommender systems deal with applications having just two types of entities, users and items. However, creation of more complex and realistic applications leads to interest in a new line of research about how to incorporate contextual information as an extra dimension into the recommendation systems (Hashemi, Clarke, Kamps, Kiseleva, & Voorhees, 2016). There are three ways of incorporating context in the recommender systems: contextual pre-filtering, contextual post-filtering, and contextual modeling (Adomavicius & Tuzhilin, 2011). As the later approach is closer to our study in this paper, we will discuss some of the related research in the contextual modeling.

In order to contextually model the context aware recommendation system, Karatzoglou, Amatriain, Baltrunas, and Oliver (2010) proposed a multiverse recommendation method based on tensor factorization, which integrates contextual information by modeling data as a User-Item-Context N-dimensional tensor instead of a traditional 2-dimensional User-Item matrix. One problem of this method is the data sparseness, which is proportional to the number of defined contexts in their method. Liu and Aberer (2013) proposed to partition the User-Item matrix by grouping ratings of similar context, which could be helpful to decrease the data sparseness. The other problem of the multiverse recommendation method is that it only works for categorical features. To overcome this problem, Rendle, Gantner, Freudenthaler, and Schmidt-Thieme (2011) proposed to use factorization machines to model contextual information. The above studies are done to model contextual information, however none of them are scalable enough to be effective for the recent exponentially growing data.

2.2. POI Recommendation Systems

There have also been many studies to solve the POI recommendation problem in both academia and industry (Guy, 2015; Zoeter, 2015). They generally try to adapt traditional recommendation algorithms to the POI recommendation problem. One line of research includes collaborative filtering and matrix factorization approaches in location-based social networks (LBSNs). Berjani and Strufe (2011) proposed regularized matrix factorization, in which they apply personalized collaborative filtering on dimensionally reduced user-POI matrices to minimize the squared regularized error. In addition to the geographical aspects, there is research on POI recommendation that in addition to the geographical dimension also includes the temporal dimension in the matrix factorization framework (Gao, Tang, Hu, & Liu, 2013; Griesner, Abdessalem, & Naacke, 2015).

Within the POI recommendation literature, there are some studies that are related to ours in the sense that they studied users' check-in behavior (Park, Hong, & Cho,

2007; Scholz, Illig, Atzmueller, & Stumme, 2014; Xiao, Zheng, Luo, & Xie, 2010; Ye, Yin, Lee, & Lee, 2011; Ying, Lu, Kuo, & Tseng, 2012; V. W. Zheng, Cao, Zheng, Xie, & Yang, 2010; V. W. Zheng, Zheng, Xie, & Yang, 2010; Y. Zheng, Zhang, Xie, & Ma, 2009; Zhuang, Mei, Hoi, Xu, & Li, 2011). As three interesting examples of these related works, V. W. Zheng, Zheng, et al. (2010) proposed collaborative location activity filtering. Particularly, they used collective factorization to recommend locations or activities to users. To this aim, they used comments having GPS data in a web-based GPS management system as a data source. Moreover, Ye et al. (2011) proposed a collaborative POI recommendation algorithm based on geographical influence. To this aim, they used users check-in activities in LBSNs. At last, Scholz et al. (2014) studied talk attendance prediction in an academic conference using a link prediction approach. To this aim, they logged talk attendance behavior using RFID tags. However, none of the above studies used both the actual users' onsite physical information interaction behaviors and users' online digital click-through behaviors.

2.3. Recommendation Systems in Cultural Heritage

Another line of related work is research on recommender systems for museum visitors. In museums, although using mobile tour guides cause negative social effects such as less interaction with visitors' fellow group members in a group visit, visitors are interested in using location-aware mobile tour guides, in which they could get information from the guide and spend more time in exhibitions (Lanir, Kuflik, Dim, Wecker, & Stock, 2013). As many museums have extensive collections of objects which makes it impossible to visit all of them in a single day, requiring visitors to be selective. Thus, personalization become one of the key topics of research in cultural heritage domain (Ardissono, Kuflik, & Petrelli, 2012).

Grieser, Baldwin, and Bird (2007) studied next exhibition recommendation problem in the museum space using visitors history. They applied Naive Bayes learning model using textual description, geospatial proximity and popularity of exhibitions. In their study, popularity baseline, which is one of our defined baseline in this paper, was reported as the most successful next exhibition recommendation model.

Bohnert, Zukerman, and Laures (2012) studied unseen exhibition recommendation using nearest-neighbor content-based filtering approach by taking visitors explicit ratings of exhibitions as inputs. They did the study using 41 museum visitors as participants. Moreover, Bartolini et al. (2016) study recommendation of diverse multimedia data across several web repositories, and arrangement of them in visiting paths. They consider location, number of persons and weather condition as context in their contextual pre-filtering system, and they did the study based on 90 users as participants.

Apart from different recommendation methods being used in the above studies in the museum domain, they are limited in term of number of participants in the experiments. In addition, none of them log and study users' onsite physical information interactions behaviors. In this paper, we log more than 21,000 users' visits of a museum in a 5 months period in operational practice, and our proposed model is based on users' both online digital and onsite physical information interaction behaviors.

In visiting a museum, recommendations can sometimes be very binary, which leads to either a satisfactory visit or a dis-satisfactory one. For example, a visitor might be in a situation of deciding a path to take from two available ones. The problem of deciding which path to target to take in museums has been addressed in (Wecker, Lanir, Kuflik, & Stock, 2011) by splitting screen of their mobile tour guide to two parts

in order to show both paths and what objects are in their way in each path. This is a critical problem that the authors address by giving information to users to decide themselves. In this paper, we address this problem by a one-shot POI recommendation system using a deep multilayer perceptron.

Closest in spirit to our work is (Hashemi & Kamps, 2017a), in which users' on-site physical behaviors in the existence of a crowd of users have been studied. They studied skip or stay behavior prediction in checking in different POIs as a classification problem. Their study is different from ours as they do not investigate on similarities between users' physical and digital behaviors. Furthermore, we study a POI ranking problem in this paper but they did research on onsite physical interaction behavior classification problem.

2.4. Internet of Things

The Internet of Things (IoT) is a network of connected physical objects, in which sensors and actuators are seamlessly embedded in physical environments, and information is shared across platforms to develop a common operating picture (Gubbi, Buyya, Marusic, & Palaniswami, 2013). The IoT was first introduced by Kevin Ashton in 1999 in supply chain management context (Ashton, 2009). Then, in the past decade, IoT applied to many applications such as health care systems (Catarinucci et al., 2015), smart cities (Zanella, Bui, Castellani, Vangelista, & Zorzi, 2014) and smart museums (Hashemi & Kamps, 2017b).

Integration of IoT in physical environments provides not only the possibility to collect information from the environment (i.e., sensing) and interact with the environment via actuation, command and control (Gubbi et al., 2013), but also the opportunity to use the collected information to provide services to users such as analytics (Strohbach, Ziekow, Gazis, & Akiva, 2015) and personalization (Evangelatos, Samarasinghe, & Rolim, 2013; Hashemi & Kamps, 2017b).

As the most relevant line of research to our study in this paper, Evangelatos et al. present a framework for creating personalized smart environments using wireless sensor networks. Similar to our proposed behavioral user model, their proposed framework can take personalized action based on some predefined profiles including information such as users' age. However, our proposed personalization model is very different from their model as we model users behavior based on their implicit interaction signals collected using sensor logs and personalize a user experience based on the user's behavior. Furthermore, their experimental results is based on just 8 users, which is much lower than the number of users in our experiments based on an operational IoT museum environment. In fact, our experimental results is based on thousands of users' onsite and online information interactions logs.

3. POI Recommendation Using Users' Behaviors

This section studies how to predict relevant POIs to the given user and context based on users' interaction behaviors, aiming to answer our first research question: *How to understand users' onsite physical behavior and create a behavioral user model that is able to effectively predict relevant unseen POIs?* To this aim, we first present how the smart museum and our collected user interaction logs look like. Then, after formally stating the POI recommendation problem, we detail our proposed behavioral user models and features extracted for training the model.



Figure 3. An interactive POI in a museum physical space and a RFID tag as a key.

3.1. POI Recommendation in Smart Museums

There is a growing interests in integration of IoT in museums aiming to provide smart services for museum visitors (Alletto et al., 2016; Ardito et al., 2018; Ceipidor et al., 2013; Chianese & Piccialli, 2014; Gribaudo, Iacono, & Levis, 2017; Mighali et al., 2015; Rao, Sharma, & Narayan, 2017; Sornalatha & Kavitha, 2017). In this study, we focus on a specific type of smart museums that aims to understanding users' information interaction behavior based on collected onsite sensor and online click-through interaction logs. In particular, we define a smart museum as:

- **Smart museum** is a museum with exhibitions that are richly and invisibly interwoven with sensors, actuators, displays, and computational elements, embedded seamlessly in museum visits, and connected through a continuous network.

The data used in this paper is based on the smart exhibition that is part of the Roman department of the Allard Pierson Museum in Amsterdam, the Netherlands. We aim at modeling users' onsite physical interaction behavior in a smart museum by training a behavioral user model based on a collected sensors' information interaction logs. To this aim, in our smart exhibition RFID tags are provided as a key to access some additional information about objects being shown in the museum. Figure 3 shows an example of how these keys are being used to unlock content at each POI. These keys are given to users at the start of the exhibition.

At the start of the museum exhibition, there is a check-in station, at which users can enter their preferences in order to personalize the content being shown in all of the POIs. These preferences are perspectives of the narratives (i.e., Rome, Egypt and Lowlands), language (i.e., English and Dutch), and the user's age range (i.e., Adults and Children). Figure 4 shows statistics of a sample of the smart museum visitors' preferences collected at the check-in station. In this sample, we exclude any user session that has missing value for any of the three collected preferences. As it is shown in Figure 4, visitors are interested in all available content perspective prepared for POIs. Furthermore, as the smart museum is in Netherlands and it is expected, visitors usually preferred Dutch over English content. Moreover, the smart museum is an archaeological museum and our collected onsite interaction logs indicates that we have more adults visitors compared to children visitors.

After checking in, users are free to put their tags on RFID readers of some or all POIs to unlock contents being shown about objects at the POIs. We are mainly in-

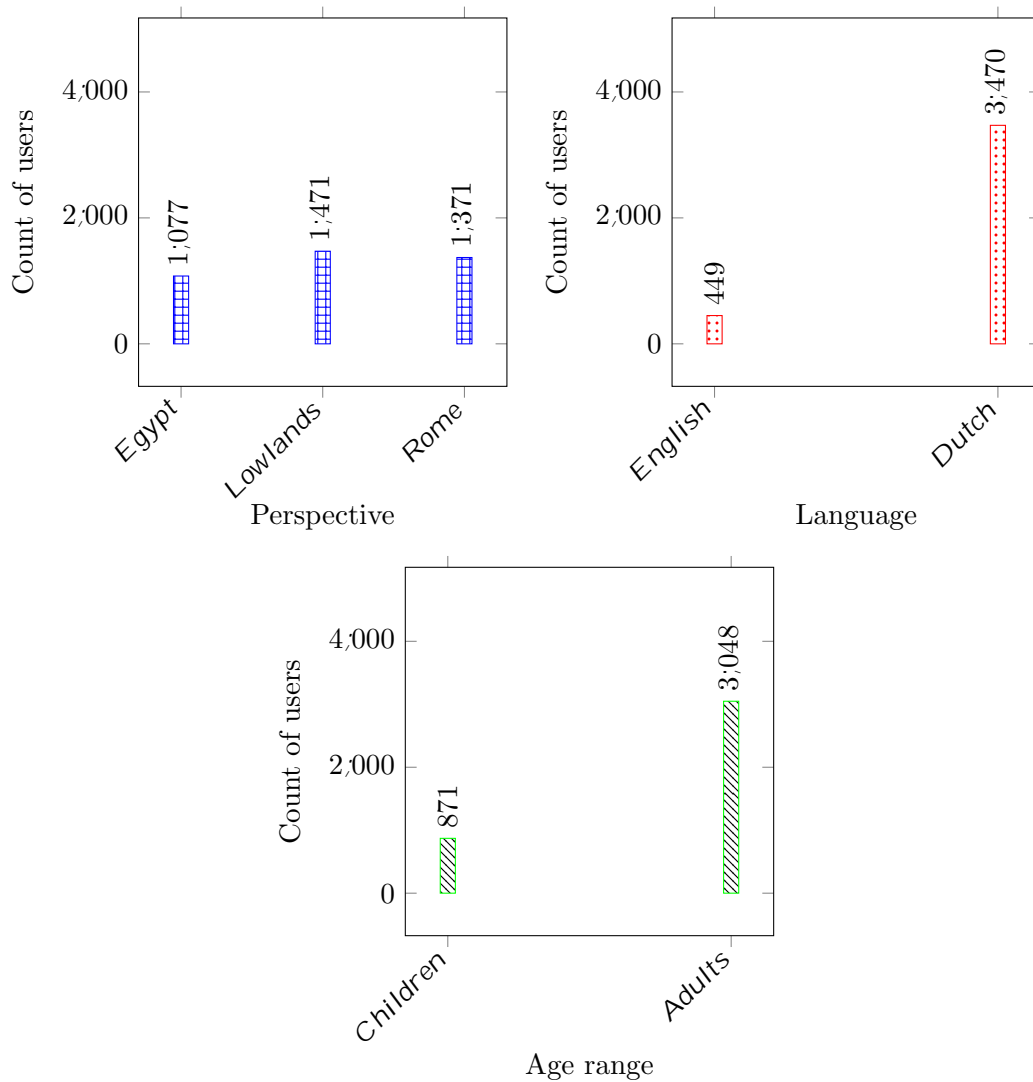


Figure 4. Distribution of onsite explicit context chosen by visitors at the check-in station.

interested in the choice, and order, of POIs visitors choose to interact with. Each POI contains three different archeological objects. Users are free to interact with POIs in any order. They can watch short movies, interact with 3D photos of POIs' objects, or read contents about objects being shown at POIs. At each POI, users are able to change the perspective of narratives and learn about objects from different perspectives. However, their visit will still be personalized based on their preference at the check-in station, and they will see narratives based on their initial choice at the next POI. At last, users might check out in a summary station, in which they might leave their name, gender, birth date and email. By leaving their email, users show their interests to receive more content about the exhibition in a post-visit scenario.

In addition to the users' onsite physical information interaction logs, we have also collected query and click-through logs of the museum search engine. Specifically, when users are in the museum website and explore the museum collection, they might search for an object by issuing a query and then clicking objects being shown in search engine result page (SERP). They might even not issue a query and just click on objects

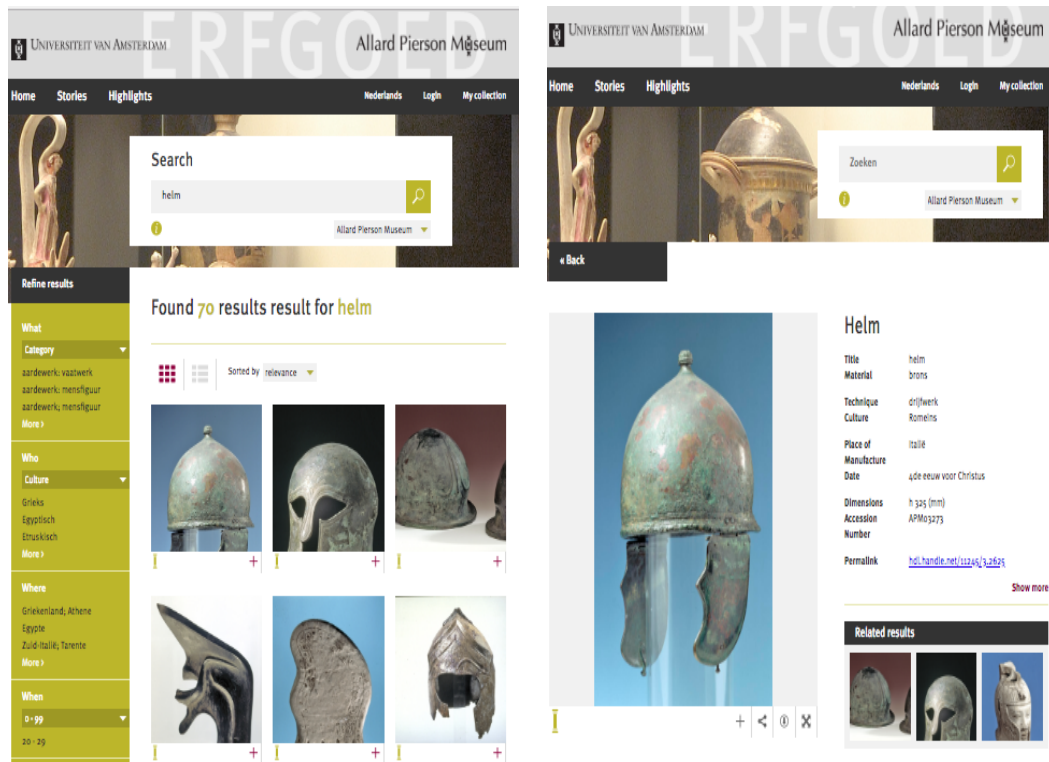


Figure 5. A museum's online collection search engine result page (left figure) and object page including related objects recommended to users based on clicking on an object presented in the search engine result page (right figure).

recommended by the museum recommender system. By clicking on objects ranked in the SERP or recommended in the museum search engine first page without issuing a query, users land on the object page, which is shown in Figure 5. In the object page, the museum recommender system recommends the most similar objects to the clicked object, which easily lead to click chaining in session. In addition, users might return to SERP and click on another object. They might also revise their query and click on objects retrieved for the given revised query. All these online users' interaction behaviors leads to click chaining that is the basis of our defined online features, which are detailed in Table 1.

There are other types of the museum search engine sessions which are not useful for collecting our online features. As all of our online features are based on users' online click-through behavior, we exclude sessions with no click in our data pre-processing. Furthermore, we filter out bot sessions in the data pre-processing.

In smart museums, there are many external factors that might have impact on users' preferences in visiting POIs. For example, a user might be interested in POIs having most popular objects in the exhibition. Furthermore, a user's check-in behavior might be affected by location of POIs presented in the museum (Hashemi, Hupperetz, et al., 2016) or even visitors' crowd in the museum (Hashemi & Kamps, 2017a). In addition to all these external factors, users' preference play a major role in their choice to visit an unseen POI after visiting a set of POIs. Users' behavioral dynamics, due to existence of all these factors, makes it very challenging to predict users' next check-in interaction after visiting a set of POIs. To address this problem, in addition to explicit context given by users at the start of an exhibition, we try to implicitly capture context by

user’s choice of visiting a set of POIs in the physical environment. In the rest of this section, we first state unseen POI recommendation problem based on a set of seen objects in a smart museum, and then we detail our proposed model to address this problem.

3.2. Problem Statement

Let $\mathbf{u} = \{u_1; u_2; \dots; u_i\} \subset U^i$ be a subset of users visited a smart environment, $\mathbf{c}_{seen} = \{c_1; c_2; \dots; c_j\} \subset C^j_{seen}$ a subset of seen or occurred contexts, and $\mathbf{p}_{seen} = \{p_1; p_2; \dots; p_k\} \subset P^k_{seen}$ a subset of seen POIs. Then, let $\mathbf{R}_{seen} \in \mathbb{R}^{i \times j \times k}_{seen}$ be a user-context-POI matrix containing i users, j seen contexts and k seen POIs. Value $r_{i;j;k} \in \mathbf{R}_{seen}$ refers to the visit frequency of user i , in context j to the POI k . In this paper, due to the fact that museum visitors rarely check in to a POI more than once, we have used binary seen or unseen values rather than considering the frequency.

Having above information about users, given a subset of unseen contexts (i.e., $\mathbf{c}_{unseen} = \{c_1; c_2; \dots; c_m\} \subset C^m_{unseen}$), and a subset of unseen POIs (i.e., $\mathbf{p}_{unseen} = \{p_1; p_2; \dots; p_n\} \subset P^n_{unseen}$), the behavioral unseen POI recommendation problem is estimation of $r_{i;m;n} \in \mathbf{R}_{unseen}$ based on users interaction behaviors with the seen POIs, in which $\mathbf{R}_{unseen} \in \mathbb{R}^{i \times m \times n}_{unseen}$ is a user-context-POI matrix containing i users, m unseen contexts and n unseen POIs.

In order to model the set-based contextual POI recommendation, we cast the context-aware recommendation problem to a binary classification problem, in which relevant POIs are labeled 1 and irrelevant ones labeled 0. In this way, we try to learn a behavioral model to predict relevant unseen POIs to the given user and context based on the user’s interaction behaviors in the context. Then, relevance probability of POIs to the user and context pairs will be used to rank the unseen POIs. To this aim, a set of features that represent users’ interaction behaviors in given contexts is defined.

3.3. Feature Set

In order to learn an effective model to rank POIs, we have extracted 18 different features. As shown in Table 1, we have classified features to three sets, namely, explicit context, onsite and online.

The *explicit context features* refer to information explicitly given by users about the context. In our study, we collected users’ gender, their preferred language, their age range and their chosen perspective of the narratives at the exhibition. Previous study on these explicit contexts (Hashemi, Hupperetz, et al., 2016) shows that users behave differently in these different contexts. For example, as it is discussed in (Hashemi, Hupperetz, et al., 2016), children tend to spend less time in front of the POI about death. Therefore, it seems a reasonable set of features to consider as explicit contexts. Furthermore, the content being shown in the exhibition at each POI is personalized, which implicitly has impact on users onsite interaction behavior.

The second group consists of *onsite features* which are a set of implicit behavioral features collected during the interactions in the smart environment. In particular, we use onsite features extracted based on user walk-through data. Specifically, f_5 is the number of seen POIs, which can be a signal of visitors’ expertise in interacting with the POIs. In addition, it can be considered as a confidence indicator of some other features’ scores like f_6 . Whereas f_6 is the content-based filtering score of POI candidate based on the profile built using the seen POIs. This content-based filtering score is

Table 1. Defined features to predict relevant unseen POIs to users after visiting a set of POIs

Feature	Category	Description
f_1	Explicit Context	Gender (e.g., Female)
f_2	Explicit Context	Language (e.g., English)
f_3	Explicit Context	Visitor age range (e.g., Adults)
f_4	Explicit Context	Chosen perspective (e.g., Roman)
f_5	Onsite	Seen POIs set size.
f_6	Onsite	Content-based relevance score of a POI candidate to a profile created using seen POIs' content that was shown onsite
f_7	Onsite	Unseen POI's PageRank in onsite visits walk-through weighted graph built based on a train set.
f_8	Onsite	Unseen POI's PageRank in onsite visits walk-through unweighted graph built based on a train set
f_9	Onsite	Unseen POI's centrality in onsite visits walk-through graph built based on a train set.
f_{10}	Onsite	Minimum distance of the seen set of POIs to the POI candidate in the onsite visits walk-through graph built based on a train set
f_{11}	Onsite	Median distance of the seen set of POIs to the POI candidate in the onsite visits walk-through graph built based on a train set
f_{12}	Onsite	Mean distance of the seen set of POIs to the POI candidate in the onsite visits walk-through graph built based on a train set
f_{13}	Online	Unseen POI's PageRank in Online click-through weighted graph built based on a train set
f_{14}	Online	Unseen POI's PageRank in Online click-through unweighted graph built based on a train set
f_{15}	Online	Unseen POI's Centrality in Online click-through graph built based on a train set
f_{16}	Online	Minimum distance of the seen set of POIs to the POI candidate in the Online click-through graph built based on a train set
f_{17}	Online	Median distance of the seen set of POIs to the POI candidate in the Online click-through graph built based on a train set
f_{18}	Online	Mean distance of the seen set of POIs to the POI candidate in the Online click-through graph built based on a train set

calculated based on the onsite POI descriptions and users’ onsite interactions. That is why it is considered as one of the onsite features in our feature classification.

In addition to f_5 and f_6 , we build users’ walk-through graph using their onsite interactions with POIs based on the train set onsite information interaction logs, and calculate the further f_7 , f_8 , f_9 , f_{10} , f_{11} and f_{12} features. Details of these features are available in Table 1. In particular, f_7 is unseen POI’s PageRank in the onsite visits walk-through weighted graph. Weight of a link from POI_a to POI_b is the number of times that visitors visited POI_b after checking in at POI_a . The main motivation behind using pagerank rather than link popularity of POIs is the fact that pagerank helps minimizing the effect of position rank bias of the POI. It is shown in (Hashemi, Hupperetz, et al., 2016) that there is a position rank bias in smart museums and it is more likely that users check in at POI1, which is the closest POI from the check-in station. This leads to high degree of both incoming and outgoing node degree for POI1. Using pagerank give less importance for incoming links from POIs with many outgoing links (e.g., POI1), which minimizes the possible bias on users’ behavior based on available external factors. On the other hand, f_9 is centrality feature that can capture popularity in the walk-through graph.

The third group consists of *online features* refers to a set of features based on on-line interaction logs based on the collection information as offered on the museum’s web site. The features are defined in a similar way as we have modeled the onsite selected POIs using the onsite users’ interactions logs. However, the feature calculation is entirely based on the prior online click-through graph of the museum search engine. As said before, we assume a cold start scenario, where no mapping between users at the smart exhibition and the online logs, hence no online prior history of the particular visitor. The online click-through graph is filtered to the objects available at onsite POIs. In this study, each onsite POI contains 3 different museum objects. We merge all the objects related to each POI as one node, and the click-through graph’s edges are aggregated from all the edges of POIs’ objects. As a result, same as onsite walk-through graph, the online click-through graph has onsite POIs as nodes. Details of these features are available in Table 1.

3.4. Learning Model

In order to learn a set-based behavioral POI recommendation model, we have implemented a logistic regression classifier and a deep neural multilayer perceptron with dropouts to estimate relevance of each POI to the given user after visiting a set of POIs. The logistic regression classifier and the deep multilayer perceptron have been trained separately based on each group of features extracted using different users’ information interaction behaviors to study which user information interaction behavior is more effective in understanding users’ preferences in their interactions with the IoT in smart environment. In the rest of this section, we will detail the logistic regression and the deep multilayer perceptron implemented for the set-based behavioral POI recommendation.

3.4.1. Logistic Regression

Logistic regression classifier is a linear classifier that transparently helps understand contribution of each feature in estimation of POIs relevancy. In fact, we would like to know which trained logistic classifier performs better and why. To this aim, we train different logistic regression classifiers based on different feature sets using different

users' interaction behaviors.

In order to learn a logistic classifier, we use variable $c \in \{0;1\}$ to show relevance of a POI to a user in a context. Specifically, $P(c = 1|u; c; \rho)$ is the relevance score of the POI ρ to the user u and the context c , in which ρ is unknown parameters learned using maximum likelihood estimation (MLE) based on the train set. Given the relevance judgments r of each POI ρ_k to a user u_i and context c_j in the train set, the likelihood L of the train set is as follows:

$$L = \prod_{i=1}^n \prod_{j=1}^m \prod_{k=1}^K P(c = 1|u_i; c_j; \rho_k)^{r_{ijk}} P(c = 0|u_i; c_j; \rho_k)^{1 - r_{ijk}}$$

in which we assume relevance judgments r are generated independently. We model $P(c = 1|u_i; c_j; \rho_k)$ by logistic function on a linear combination of features created based on each specific group of users' information interaction behaviors. Then, we optimize the unknown parameters ρ_k by maximizing the following log likelihood function:

$$\begin{aligned} &= \underset{\rho_k}{\operatorname{argmax}} \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^K r_{ijk} \log P(c = 1|u_i; c_j; \rho_k) \\ &\quad + (1 - r_{ijk}) \log P(c = 0|u_i; c_j; \rho_k) \end{aligned}$$

In order to turn the logistic classifier scores to probabilities, we have used the softmax function:

$$S(y_i) = \frac{e^{y_i}}{\sum_j e^{y_j}}$$

in which y_i is the logistic classifier score, and $S(y_i)$ is the output relevance probability of our behavioral POI recommendation model. At last, we rank unseen POIs based on the logistic classifier output probability of POIs' relevancy being estimated based on features created using interaction behaviors of a given user in a context.

3.4.2. Deep Neural Multilayer Perceptron

In this subsection, we investigate on a deep neural multilayer perceptron (MLP) by an aim of improving effectiveness of the POI recommendation to be used in critical one-shot POI recommendation applications. The motivation behind the critical one-shot POI recommendation is that an irrelevant recommendation sometimes has a very negative effect in users' experience in a way that they might be incorrectly guided to an uninteresting department of a museum that leads to a dissatisfied experience. In this model, for each user in a context, our main goal is to recommend a POI which is highly relevant to them. In the one-shot POI recommendation, we do not care about relevant POIs retrieved after rank 1. In the rest of this section, we detail our deep multilayer perceptron with an aim of improving effectiveness of POI recommendation to be used for the critical one-shot POI recommendation problem.

In order to learn a set based behavioral POI recommendation and learn users' on-site complicated physical behaviors, we have used a deep MLP neural network with 3 hidden layers having 326 units. To learn an effective model and overcome overfitting problem, we have used a dropout feedforward neural network. Let $l \in \{1;2;3\}$ be the

index of the hidden layers of the network. Let $z^{(l)}$ be the vector of input to layer l and $y^{(l)}$ be the vector of outputs from layer l . The dropout neural network is modelled as follows for any hidden unit i and $l \in \{0; 1; 2\}$ (Hinton, Srivastava, Krizhevsky, Sutskever, & Salakhutdinov, 2012; Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014):

$$\begin{aligned} r^{(l)} &\sim \text{Bernoulli}(\rho); \\ \tilde{y}^{(l)} &= r^{(l)} * y^{(l)}; \\ z_i^{(l+1)} &= w_i^{(l+1)} \tilde{y}^{(l)} + b_i^{(l+1)}; \\ y_i^{(l+1)} &= f(z_i^{(l+1)}); \end{aligned}$$

where $r^{(l)}$ denotes a vector of independent Bernoulli random variables having probability ρ of being 1, $\tilde{y}^{(l)}$ is thinned outputs created by multiplying a sample of $r^{(l)}$ vector by outputs of layer l (i.e., $y^{(l)}$) and used as input for the next layer $l + 1$, $w^{(l)}$ and $b^{(l)}$ are weights and biases at layer l , and f is an activation function, which is rectified linear units (ReLU) in our setup. This process is done at each layer.

Following prior research in neural network domain, we have used $\rho = 0.5$ in our dropout network. This value is reported as a close to optimal value for a wide range of networks in different applications (Srivastava et al., 2014).

In the learning phase, the derivatives of the loss function are backpropagated through the dropout network. The dropout network is trained using the stochastic gradient descent (SGD) algorithm with mini batches, which is widely used algorithm for training neural networks. The learning rates are adjusted based on adaptive gradient algorithm (AdaGrad) (Duchi, Hazan, & Singer, 2011). In the test phase, the sub-network is used without dropout, but the weights are scaled as $W_{test}^{(l)} = \rho W^{(l)}$.

For the classification purpose and having probabilities as outputs, we have used Logistic classifier in the last layer. The logistic classifier in the last layer is trained same as the logistic regression classifier being discussed in previous subsection. The only difference is that, in the logistic classifier being used in the last layer, we model $P(c = 1 | u_i; c_j; p_k)$ by logistic function on a linear combination of inputs from the last hidden layer units' outputs. At last, the final relevance probability of $P(c = 1 | u_i; c_j; p_k)$ is used to rank unseen POIs based on features created using interaction behaviors of a given user in a context.

4. Experimental Setup

In this section, we describe our experimental setup. We first describe the data set used in this paper, and second detail the evaluation methodology used in this study.

4.1. Dataset

The dataset of this study is based on onsite physical and online digital interaction logs collected at an archeological museum. Onsite physical interaction logs are collected using sensors available in the museum, and the online digital interaction logs are based on click-through behavior of users.

In this paper, 5 months onsite physical interaction logs of the museum with more

Table 2. An example of records created for the test collection using a user session. The judgments are based on seen POI set-size 2 and 3

Query context	Seen POI set	Candidate	Relevance
c_1	$\langle \text{POI}_1; \text{POI}_2 \rangle$	POI_3	0
c_1	$\langle \text{POI}_1; \text{POI}_2 \rangle$	POI_4	1
c_1	$\langle \text{POI}_1; \text{POI}_2 \rangle$	POI_5	0
c_1	$\langle \text{POI}_1; \text{POI}_2 \rangle$	POI_6	0
c_1	$\langle \text{POI}_1; \text{POI}_2 \rangle$	POI_7	1
c_1	$\langle \text{POI}_1; \text{POI}_2 \rangle$	POI_8	0
c_1	$\langle \text{POI}_1; \text{POI}_2; \text{POI}_4 \rangle$	POI_3	0
c_1	$\langle \text{POI}_1; \text{POI}_2; \text{POI}_4 \rangle$	POI_5	0
c_1	$\langle \text{POI}_1; \text{POI}_2; \text{POI}_4 \rangle$	POI_6	0
c_1	$\langle \text{POI}_1; \text{POI}_2; \text{POI}_4 \rangle$	POI_7	1
c_1	$\langle \text{POI}_1; \text{POI}_2; \text{POI}_4 \rangle$	POI_8	0

than 21,000 sessions is used, which leads to 3,925 high-quality onsite sessions to be used for evaluation purposes.

The online features, detailed in Table 1, have been extracted based on 18,001 high-quality sessions created based on a common time-oriented session identification approach in search engines using 30 minutes inactivity time as session cut-off boundary (Eickhoff, Teevan, White, & Dumais, 2014; Shokouhi, Ozertem, & Craswell, 2016). The main assumption is that a long period of inactivity between a user’s activities indicates the user is probably no longer active, which leads to ending the session.

4.2. Evaluation Methodology

In our collected onsite information interaction logs, about 16,000 out of 21,000 sessions either did not have any interactions with POIs or they did not check out at the summary station, and about 1,000 of them had interactions with all the POIs. In order to avoid bias over users who are interested in visiting all or none of the POIs at the museum, we exclude all sessions have checked in at all or none of the POIs at the exhibition. As a result of this preprocessing step, 3,925 out of 21,000 high-quality onsite information interaction sessions remain for creating the test collection.

Considering the walk-through graph, for each user in a session and at each checked-in POI during their visit, we created a test collection using the seen set of POIs, the user and the explicit contexts as the query and the unseen POIs as the candidates, for which we have judgments based on the user’s session. Basically, we know which POI candidates are visited by the user and consider them as relevant POIs. The rest of the POIs are considered as irrelevant POIs.

Doing the above procedure in building the test collection leads to create a contextual set-based POI recommendation test collection having 1,083,623 judgments. Table 2 shows an example of records created using a user session. To test our proposed model, in order to avoid overfitting, we have done five-fold cross-validation, in which for each fold as a test set, three out of the four remained folds randomly sampled and used as a train set, and the remained fold used as a validation set. We repeat the process for all the five folds and report the average of the evaluation metrics.

4.3. Evaluation Metrics

For the evaluation of the defined set-based behavioral POI recommendation task, we cast the problem to a ranking task and use mean reciprocal-rank (*MRR*), mean average precision (*MAP*) and R-precision (*R-Prec*) as metrics that are effective to evaluate proposed models. Moreover, in order to evaluate the one-shot POI recommendation systems, we use precision at rank 1 (*P@1*) as an evaluation metric.

The *MRR* is the average of the reciprocal ranks of the first relevant result for a set of queries Q as:

$$MRR = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{rank_i};$$

In our experiments, Q is a set of 1,083,623 queries (user and context pairs). In *MRR*, $rank_i$ represents rank of first relevant POI for a given pair of user and context. Precision at rank n (i.e., $p@n$) is used in number of evaluation metrics in this study, which is defined as follows:

$$p@n = \frac{\# \text{ relevant POIs in top } n \text{ results}}{n};$$

where n is the rank. For a single query, *AP* is defined as the average of the $p@n$ values for all relevant POIs as:

$$AP = \frac{\sum_{n=1}^N p@n \times rel(n)}{R};$$

in which N is the number of retrieved POIs candidates, and $rel(n)$ is a binary function indicating the relevance of a POI to a given user and context pair at a given rank. A POI is relevant to a user and context pair, if the user checks in at the POI at that visit. *MAP* is the mean value of the *APs* computed for all queries. *R-Prec* is precision at rank R where R is the number of relevant candidates for the given query. At last, *P@1* is the precision at rank 1.

4.4. Baselines

In this section, we detail the baselines created for the evaluation purposes.

4.4.1. Popularity

The popularity based recommendation ranks POIs candidates according to their popularity scores. According to previous evaluation studies in recommender systems such as (Herlocker, Konstan, Terveen, & Riedl, 2004), systems recommending very popular items can guarantee that users will like most of the recommended items. Moreover, the popularity baseline is usually used in evaluation of personalized recommendation systems and it is informed as a competitive baseline (Lucchese, Perego, Silvestri, Vahabi, & Venturini, 2012).

In this paper, the popularity is computed as the number of users who checked in at each POI. Therefore, regardless of what POI has been already seen by a user, the

popularity baseline recommends the most popular POIs according to other users who checked in at the POIs before.

4.4.2. Bias-Based Filtering

In both physical and digital worlds, external factors has impact on users' behavior with information systems (Hashemi, Hupperetz, et al., 2016; Hashemi & Kamps, 2017a, 2017b). As a result, assuming existence of the same external factors in the physical smart environments, we could take advantage of them and predict the next POI based on users' status in the environment. Although the bias-based filtering baseline could be hard-to-beat, it would not be a very useful recommender system in practice. Such a baseline is not based on users' interests and their profile. They are just predicting users next move using biases and external factors in the environment.

As Hashemi, Hupperetz, et al. (2016) discussed, there are some biases in onsite user information interaction logs. They introduces the walk-through position-bias that shows users tend to visit POIs one after the other from check-in to check-out stations. They also observed time-rank bias that indicates users tend to spend less time at the end of exhibitions. Considering these two biases, the probability of checking in at a POI is proportional to the distance from the Check-out station. Therefore, in all experiments of this paper, the bias-based baseline ranks POIs based on their distance from the check-out station.

4.4.3. Content-Based Filtering

As descriptions of POIs in museums are well curated, they are an informative source of information that makes content-based filtering as an effective baseline in this domain. In this study, each POI contains three museum objects with reach descriptions. In order to build a content-based filtering model, we build a profile of each user after visiting a set of POIs using Language Modeling framework. Each profile's language model is based on all seen objects of \mathbf{p}_{seen} .

Since we have profiles of users at each context based on their seen POIs, KL-Divergence of each unseen POI's language model and the profile is considered as content-based filtering scores for ranking unseen POIs.

5. Experimental Results

In this section, we provide answer to the research questions stated in the introduction section.

5.1. POI Recommendation using Users' Information Interaction Behaviors

This section answer our second research question: *How strong are different users' interaction behaviors with IoT in understanding users' preferences?*

To this aim, we have used each of the three groups of features extracted based on each information interaction behaviors to train a POI recommendation system. Specifically, we have trained three different logistic regression classifiers, which are trained based on: 1) the explicit context features (i.e., Logistic Regression-Explicit Context) 2) the onsite features (i.e., Logistic Regression-Onsite) and 3) the online

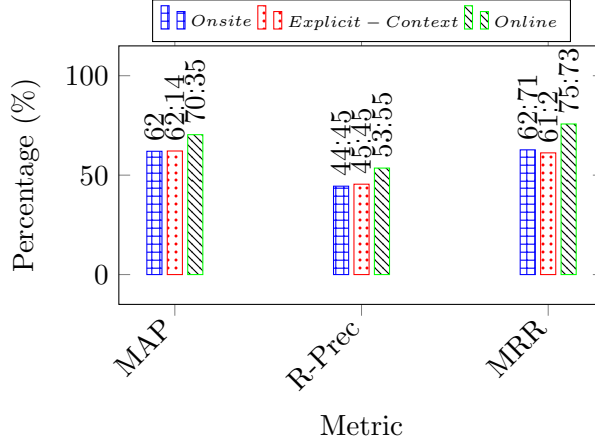


Figure 6. Effectiveness of different types of users' interaction behavior in understanding their onsite preferences.

features (i.e., Logistic Regression-Online).

In the rest of this subsection, we first investigate whether users' online digital interaction behaviors are similar to the users' onsite physical behavior. Then, we detail relative importance of each feature extracted based on features' weights being learned by logistic regression classifiers using each type of users' interaction behaviors with an aim of understanding users' behaviors.

5.1.1. Onsite Physical Behavior vs. Online Digital Behavior

We first look at the question: *Are online digital behaviors similar to onsite physical behaviors? Does understanding online digital users' information interaction behaviors have a positive effect in learning a model to predict unseen relevant POIs and complete users' personalized onsite visits?*

In order to answer this research question, we compare POI recommendation systems trained based on each type of interaction behavior. As shown in Figure 6, the POI recommendation system trained based on users' online digital interaction behavior is not only as good as the other POI recommendation systems being trained based on either explicit context or onsite interaction behaviors, but also is performing better than them in terms of all common tested information retrieval metrics.

This experiment indicates that availability of the considerable amount of online interaction logs in comparison to onsite interaction logs leads to training an effective onsite POI recommendation system based on users' online digital interaction behaviors. As we achieve an effective onsite POI recommendation system based on users' online digital interaction behaviors, we conclude that there is a similarity between onsite physical and online digital information interaction behaviors.

5.1.2. Features Relative Importance in Understanding Users' Interaction Behaviors

We now look at the question: *What are the relative importance of each feature extracted based on different users' interaction behaviors in effectiveness of POI recommendation systems?*

To this aim, we normalize features' weights being learned in each logistic regression classifier trained for each group of features separately. Then, average of the normalized features' weights over the 5-fold cross-validation are reported and compared in

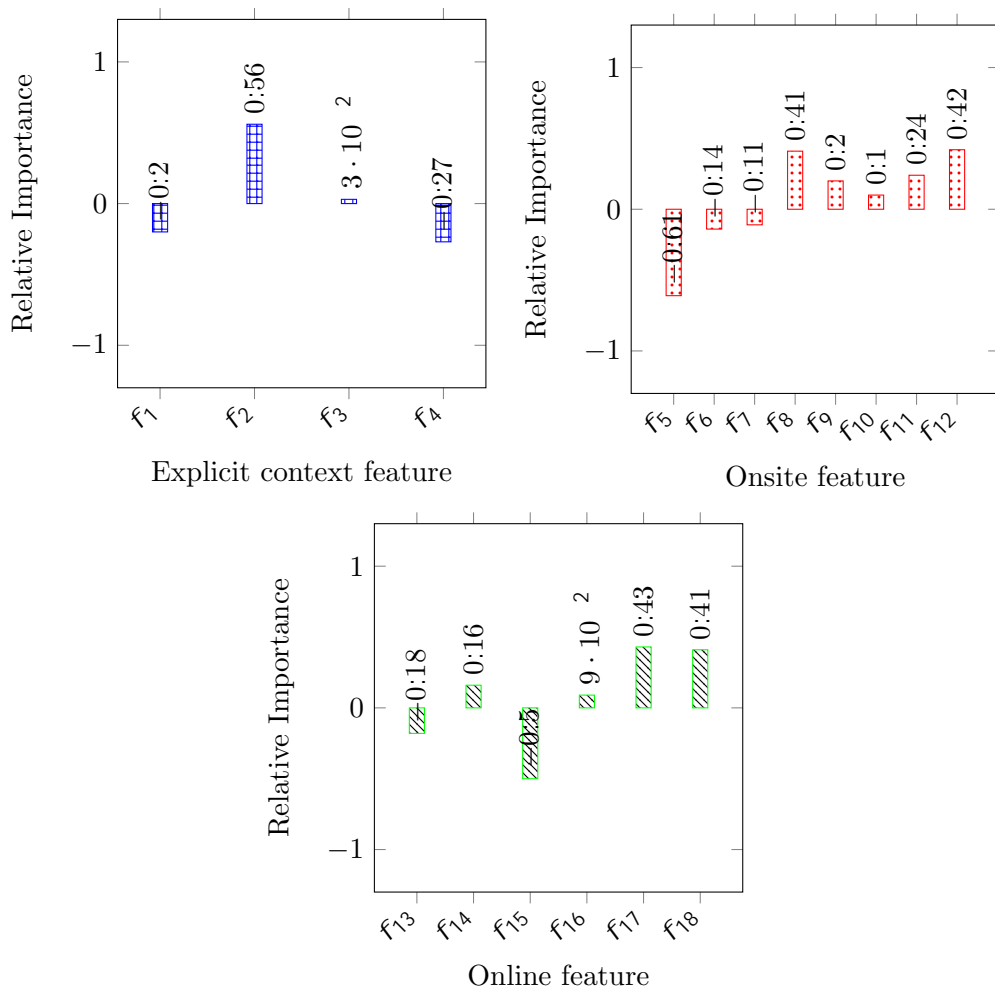


Figure 7. Features' relative importance in POI recommendation trained based on each group of users' information interaction behaviors.

Table 3. Set-based one-shot POI recommendation baselines effectiveness

Run	P@1	MRR
<i>Content-based Filtering</i>	57.45	75.68
<i>Popularity</i>	60.86	77.67
<i>Bias-Based Filtering</i>	61.57	77.71

Table 4. Set-based one-shot POI recommendation effectiveness comparison between the *Deep MLP-Online* and the best baseline. * indicates the improvement is statistically significant ($p < 0.05$)

Run	P@1	MRR
<i>Bias-Based Filtering</i>	61.57	77.71
<i>Logistic Regression-Online</i>	56.97	75.73
<i>Deep MLP-Online</i>	75.81 (23.12%*)	86.39 (11.17%*)

Figure 7.

As it is shown in Figure 7, among the explicit context interaction, the chosen language (i.e., f_2) at the start of museum visits is relatively more important in comparison to other explicit context based features. Furthermore, mean distance of the seen POIs to a POI candidate in the onsite visits’ walk-through graph (i.e., f_{12}) has relatively more importance in comparison to other onsite interaction behavior based features. Regarding the online interaction behaviors, median distance of the seen set of POIs to the given candidate in the online click-through graph (i.e., f_{17}) is relatively more important than other online features in the effectiveness of the POI recommendation systems.

5.2. One-Shot POI Recommendation Using Users’ Interaction Behaviors

This section answer our third research question: *How effective is behavioral POI recommendation system in one-shot POI recommendation problem?* To this aim, we first study effectiveness of the discussed baselines in one-shot POI recommendation problem. Table 3 shows effectiveness of the baselines in terms of $P@1$ and MRR metrics. Experimental results indicate that the Bias-based filtering baseline performs better than the other baselines in terms of both one-shot POI recommendation evaluation metrics. One possible explanation of this could be that users’ interaction behaviors is highly affected by external factors in physical environments (Hashemi, Hupperetz, et al., 2016; Hashemi & Kamps, 2017a), which leads to more predictable user behavior in the existence of those external factors. Thus, the bias-based filtering baseline is even performing slightly better than the popularity baseline, which is a hard-to-beat baseline according to previous studies in recommendation systems in cultural heritage (Lucchese et al., 2012).

In order to evaluate effectiveness of our proposed one-shot POI recommendation model, we study effectiveness of the implemented deep multilayer perceptron in one-shot onsite POI recommendation problem in comparison to the best performed baselines as well as the logistic regression POI recommendation system. Table 4 shows

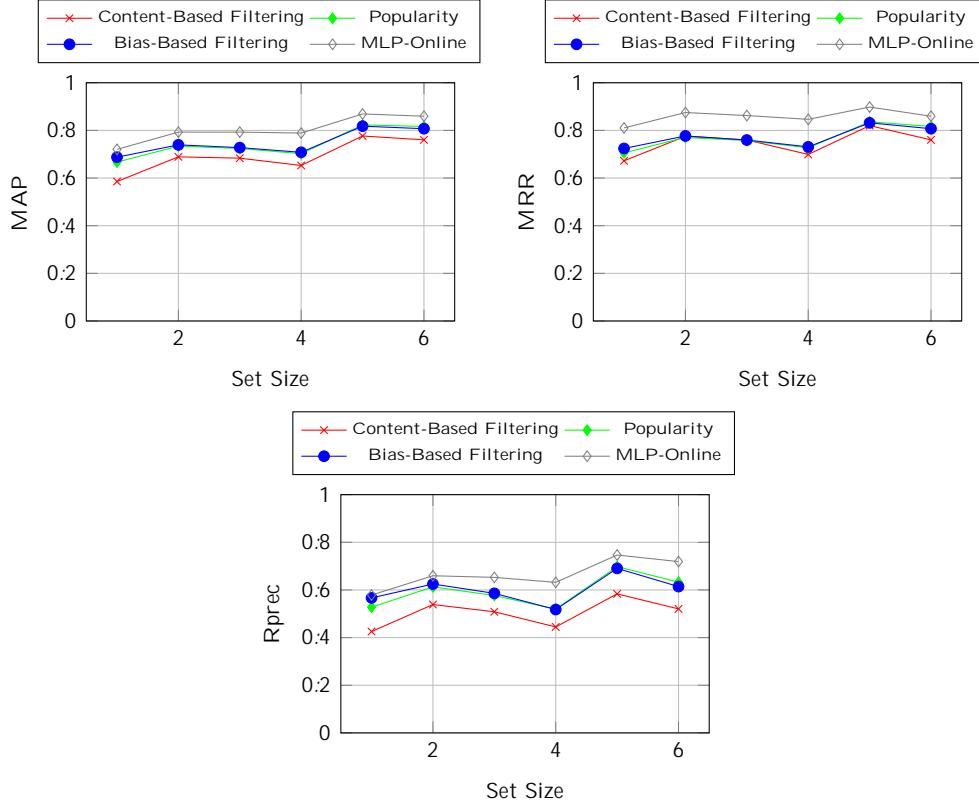


Figure 8. Effects of seen POIs set size on the performance of the best proposed model and baselines.

performance of the best deep multilayer perceptron (i.e., Deep MLP) and logistic regression classifiers, trained based on online digital interaction behaviors, in terms of $P@1$ and MRR .

In this experiment, we just focus on the results based on $P@1$ and MRR as in one-shot POI recommendation problem, we just care about the first ranked unseen recommended object. Thus, $P@1$ is the main metric in evaluation of this problem. In the evaluation of this experiment, we have also used the MRR metric as a representative of the early precision based metrics.

As it is shown in Table 4, the deep MLP significantly improves the best competitive baseline (i.e., Bias-Based Filtering) in one-shot POI recommendation. In particular, the deep MLP has 23.12% improvement over the bias-based filtering baseline in terms of $P@1$, which is the metric that measures as closely as possible the one-shot POI recommendation performance. This experimental result shows that our proposed deep MLP one-shot POI recommendation system results in very high precision, suggesting it’s practical use to created an enhanced personalized experience for this critical application.

5.3. Impact of Seen Set Size

This section answer our research question: *What is the effect of given seen POIs set-size in the unseen POI recommendation performance?*

In this experiment, we analyze impact of different seen POIs set-size in the effectiveness of the final POI recommendations. As it is shown in Figure 8, overall performance

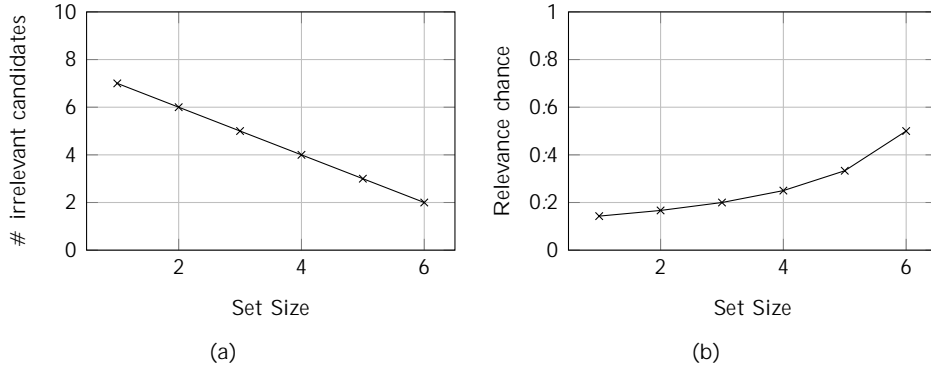


Figure 9. Impact of seen POIs set size on one-shot unseen POIs recommendation based on number of irrelevant candidates in contrast to just one relevant POI (left figure) and relevancy chance of a random recommended unseen POI (right figure).

of the recommendations are improved while users interact more with the POIs and see more POIs. However, there are some biases in the users’ onsite information interaction logs that add some noises in the observed patterns based on the seen POIs set-size.

For example, due to the observed position-rank bias in users’ onsite behavior, POI_4 has a higher chance of being the fourth seen POIs in users’ visits. It seems the POI_4 ’s location in the exhibition is the start location of a strong position-bias in which users tend to visit POIs one after the other, except when there is a crowd of visitors in front of the next POI. This makes it more difficult to understand users’ preferences in checking in at unseen POIs. This may explain why there is a slightly decrease in recommendations performance at seen POI set-size 4.

In real applications, we may always have external factors like the topology of POIs in the physical space contributing in users’ behavior, which in this case, decreases recommendations performance at four seen POIs. However, in an experimental environment that all external factors and biases are avoided, we can improve recommendations effectiveness by having more seen POIs and creating a richer profiles.

Experimental result shown in Figure 8 indicates that the effectiveness of our best proposed POI recommendation model (MLP-Online) is improved by increasing number of seen POIs in sessions. However, the improvement is not just due to obtaining more history about the user profile. Specifically, in one-shot POI recommendation problem, according to the number of available candidates in different seen POIs set-size, the one-shot POI recommendation problem becomes easier when a smaller number of unseen POIs remains, compared to the start of exhibition’s visit.

Figure 9 shows what is chance of recommending relevant POI in one-shot POI recommendation is by randomly recommending a POI at each seen POIs set-size. Specifically, when seen POIs set-size of a user visit is equal to one, one-shot POI recommendation system has seven different POI candidates to recommend in our experiment. As a result, by just randomly recommending a POI, it would have $1=7 \sim 14\%$ chance of recommending a relevant POI to the user. On the other hand, if a user visited six POIs and has two unseen POIs in their visit, we would have $1=2 \sim 50\%$ chance of recommending a relevant POI to the user by randomly recommending a POI.

As it is discussed in this experiment, our proposed model based on online features is much less affected by the available biases in the users’ onsite information interaction logs in comparison to all the other models. This experiment shows that the proposed model is performing better than all the baselines at any seen set-size. In fact, although one-shot POI recommendation problem is relatively more difficult when a user’s seen

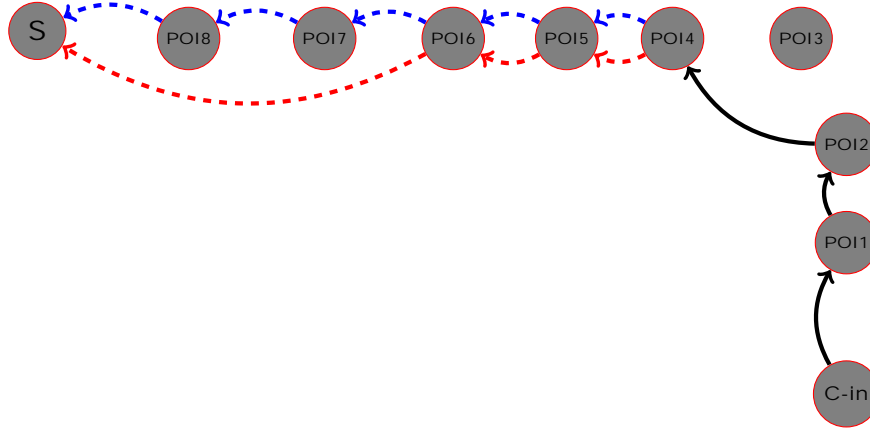


Figure 10. Users' unseen POI routes after visiting a set of POIs, namely, POI1, POI2, and POI4 by skipping the POI3. The figure demonstrates two most popular unseen POI routes based on a real traffic in a smart museum. Each of them shown by a different color, and the black edges are the ones walked by all the three visitors. *C-in* is the check-in station and the *S* is the check-out station.

POI set-size is low and have relatively higher number of candidates compared to later stage of their visit, the improvement is even higher in lower seen set-sizes. One possible explanation of this is that as the MLP-online trains the one-shot POI recommendation model based on a larger number of hyper-parameters compared to baselines, it could be able to have a greater improvement over baselines when the problem is harder to address. In the next section, we discuss what would be future directions of our study in POI recommendation in smart environments.

6. Future Directions

As shown in this paper, we have achieved a high performance for next POI recommendation problem using our proposed model. This one step recommendation problem is a key application for museum exhibition navigation, or more generally next step recommendation in smart environments, but there are other interesting applications that suggest themselves. In particular, can we recommend a whole route which may require additional aspects such as considering length or diversity, that are not captured by the one step recommendation problem. In future work, we plan to study the problem of route prediction in smart environments based on seen POIs profile logged by onsite sensors.

Let us discuss an illustrative example. In a sample of the onsite sensor logs of the smart museum being studied in this paper, we have got 136 visitors who have checked-in POI_1 , POI_2 and POI_4 but decided to skip interacting with POI_3 . At this point, it would be interesting to recommend a personalized route to users. According to our observation, users behave differently in checking-in the remained unseen POIs, namely,

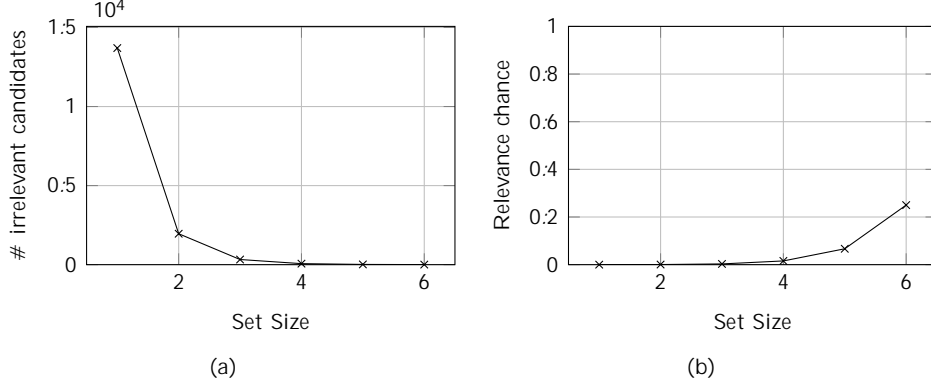


Figure 11. Impact of seen POIs set size on unseen POIs route recommendation based on number of irrelevant candidates in contrast to just one relevant route (left figure) and relevancy chance of a random recommended unseen POIs route (right figure).

POI_5 , POI_6 , POI_7 , and POI_8 . In particular, 18% (24 out of 136) of the sampled users chose to visit all the remained POIs one after the other (blue dashed lines in Figure 10), which is the most popular route. The second popular route is checking-in POI_5 and POI_6 but skip interacting with POI_7 and POI_8 (red dashed lines in Figure 10), which was based on 12 % (16 out of 136) of the sampled users’ interactions.

As it is shown above, visitors have different preferences in checking-in different POIs. Thus, understanding users’ onsite interaction behavior and recommending the best route to take in smart environments is a challenging problem to study. We do not discuss ideas on how to model users’ behavior to predict unseen objects’ route, however, we have observed different behaviors based on some explicit preferences that were given by visitors. For example, among the 24 visitors who decided to check-in all the remained POIs of the above example, 10 out of 24 were interested in narratives from “low lands” perspective in contrast to 7 out of 24 who were interested in narratives from “Rome” perspective. The rest were interested in narratives from “Egypt” perspective. As we have observed for the POI recommendation in smart environments problem, using explicit-context, onsite and online features leads to effective POI recommendation models. Thus, in the future work, using the mentioned features might be also a reasonable features to start for the unseen route recommendation problem in smart environments.

As it is discussed in previous section, the seen POI set-size has a direct impact on number of unseen objects in smart museums, which has effect on difficulty of predicting relevant POI in the one-shot POI recommendation problem. Similarly, we have studied impact of seen POI set-size on number of candidate routes, in which just one of the routes is the relevant one. Figure 11 shows number of candidate routes at each seen POIs set-size, which is calculated based on the following equation:

$$N_{rc} = \sum_{k=1}^n \frac{n!}{(n-k)!};$$

where k is the size of sequence of unseen predicted POIs, n is the number of unseen POIs in a user’s session, and N_{cr} is the total number of route candidates having from one to n route length (number of POIs in the recommended route). As it is shown in Figure 11, due to the number of irrelevant routes available for each relevant unseen

POIs route, the unseen POI route recommendation is a more challenging problem to address compared to the next POI recommendation problem. We leave investigation on this problem in smart environment to a future work.

One could easily think of further extensions of this, using the same kind of techniques to address related problems emphasizing different aspects. Of particular interest is to look the social aspects of smart exhibition visits, and ways to bring social aspects into the digital realm. A specific interesting problem to tackle here is recommending the most similar visitors, rather than items or object, in the smart environment. This could be a great strategy to bring the social aspect to museum visits. As it is discussed in (Lanir et al., 2013), using mobile tour guides has negative social effects such as less interaction with visitors’ fellow group members in a group visit. However, recommending similar users in a museum who are most likely take a same route and visit same objects, we can motivate individual users to create a group whose members have similar preference. In this way, we could have a positive impact on social aspect of museum visits, by showing the steps of prior, like-minded visitors, and bring the museum and the digital alive.

7. Discussion And Conclusions

The main focus of this paper is the study of how to build a behavioral user model for the set-based POI recommendation problem using users’ both onsite and online information interaction behaviors. Our study on the strength of using each type of users’ interaction behaviors with IoT in understanding users’ onsite information interaction preferences shows that POI recommendation systems trained using features extracted from a combination of both onsite physical and online digital information interaction behaviors (i.e., online features) performs better than the ones trained by explicitly given context or onsite information interaction behavior. Therefore, we conclude that there is a similarity between onsite physical and online digital interaction preferences that causes an improvement on the onsite POI recommendation effectiveness.

Furthermore, we have studied the critical one-shot POI recommendation problem. According to our analysis, the learned models based on just basic explicit given contexts or onsite users’ behaviors do not improve the hard-to-beat defined baselines (i.e., popularity and bias-based filtering). However, using a deep multilayer perceptron based on features extracted by online interaction behaviors leads to a significant improvement over the best baseline in all the defined evaluation metrics. Specifically, it has a statistically significant improvement over all baselines with 23% improvement in term of $p@1$ and 11% improvement in term of MRR . Therefore, our proposed approach is very effective in critical one-shot POI recommendation. Furthermore, we have studied impact of seen objects set size on the performance of the proposed POI recommendation systems. According to our experiment, the recommendation performance is generally increased proportional to the seen object set size. Although external factors have impact on users’ behavior at seen set size four in the exhibition, our proposed deep MLP model based on online features is less sensitive to the external factors and performs better than other models and baselines at all seen objects set sizes. Our proposed MLP approach achieves 83% precision at rank 1 on the critical one-shot POI recommendation problem, realizing the high accuracy needed for fruitful deployment in practical situations. The proposed behavioral user model is generic and can be widely used in any environment with an integrated Internet of Things (IoT) infrastructure. Specifically, in the Cultural Heritage domain, the IoT applications hold

the promise to provide a more interactive and multisensory experiences for visitors, and is expected to be integrated into museum practice in the next years (Freeman et al., 2016; Johnson, Adams Becker, Estrada, & Freeman, 2015). Our proposed model exploits online features hence is only applicable in cases where an online search engine with the similar objects or content related to the POIs is available for extracting the online features. Although many museums and organizations have a website with a search engine on their collection, it may not be the case in other applications in different types of smart environments.

Our general conclusion is that it is possible to fruitfully combine information interactions in the online and physical world for effective recommendation in smart environments, thereby effectively blending real-world and online behavior in principled ways. This is an attractive direction, as IoT data is typically far more sparse than online data due to physical or geographical constraints on users requiring to be physically in the smart space. In future work, we are going to increase number of POIs in the museum and see how effective is the proposed behavioral set-based POI recommendation model for a bigger dataset. We will also study the unseen POI route recommendation problem in smart environments, which is a more challenging problem to address compared to the next POI recommendation. In order to bring social aspect to users' visit in a smart museum, we will study similar person recommendation problem with an aim of creating groups with a similar preference, which might lead to similar unseen POIs routes in their visit.

Acknowledgments

This research is funded in part by the European Community's FP7 (project meSch, grant # 600851), see <http://www.mesch-project.eu/>. We especially thank the Allard Pierson Museum of Archaeology, <http://www.allardpiersonmuseum.nl/en>, its visitors, and our collaborators Wim Hupperetz and Merel van der Vaart for the chance to run extensive experiments using the museum as a lab for innovative applications. We also thank Dominique Rau and Thomas Kubitzka from the University of Stuttgart's HCI Lab for their help in developing the smart museum application.

References

- Adomavicius, G., & Tuzhilin, A. (2011). Context-aware recommender systems. In *Recommender systems handbook* (pp. 217–253). Springer.
- Alletto, S., Cucchiara, R., Fiore, G. D., Mainetti, L., Mighali, V., Patrono, L., & Serra, G. (2016). An indoor location-aware system for an IoT-based smart museum. *IEEE Internet of Things Journal*, 3(2), 244–253.
- Ardissono, L., Kuflik, T., & Petrelli, D. (2012). Personalization in cultural heritage: the road travelled and the one ahead. *User Modeling and User-Adapted Interaction*, 22(1), 73–99.
- Ardito, C., Buono, P., Costabile, M. F., Desolda, G., Lanzilotti, R., Matera, M., & Piccinno, A. (2018). Towards enabling cultural-heritage experts to create customizable visit experiences. In *Proceedings of the 2018 AVI-CH workshop on advanced visual interfaces for cultural heritage* (Vol. 2091).
- Ashton, K. (2009). That internet of things thing. *RFID Journal*, 22(7), 97–114.
- Atzori, L., Iera, A., & Morabito, G. (2010). The internet of things: A survey. *Computer networks*, 54(15), 2787–2805.

- Barnaghi, P., Wang, W., Henson, C., & Taylor, K. (2012). Semantics for the internet of things: early progress and back to the future. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 8(1), 1–21.
- Bartolini, I., Moscato, V., Pensa, R. G., Penta, A., Picariello, A., Sansone, C., & Sapino, M. L. (2016). Recommending multimedia visiting paths in cultural heritage applications. *Multimedia Tools and Applications*, 75(7), 3813–3842.
- Berjani, B., & Strufe, T. (2011). A recommendation system for spots in location-based online social networks. In *Proceedings of the 4th workshop on social network systems* (pp. 4:1–4:6). ACM.
- Bohnert, F., Zukerman, I., & Laures, J. (2012). GECKOmmender: Personalised theme and tour recommendations for museums. In *International conference on user modeling, adaptation, and personalization* (pp. 26–37). Springer.
- Catarinucci, L., de Donno, D., Mainetti, L., Palano, L., Patrono, L., Stefanizzi, M. L., & Tarricone, L. (2015). An IoT-aware architecture for smart healthcare systems. *IEEE Internet of Things Journal*, 2(6), 515–526.
- Ceipidor, U. B., Medaglia, C. M., Volpi, V., Moroni, A., Sposato, S., Carboni, M., & Caridi, A. (2013). NFC technology applied to touristic-cultural field: A case study on an italian museum. In *2013 5th international workshop on near field communication (NFC)* (p. 1-6).
- Chianese, A., & Piccialli, F. (2014). Designing a smart museum: When cultural heritage joins IoT. In *2014 eighth international conference on next generation mobile apps, services and technologies* (p. 300-306).
- Chuklin, A., Markov, I., & Rijke, M. d. (2015). Click models for web search. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 7(3), 1–115.
- Duchi, J., Hazan, E., & Singer, Y. (2011). Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12(Jul), 2121–2159.
- Eickhoff, C., Teevan, J., White, R., & Dumais, S. (2014). Lessons from the journey: A query log analysis of within-session learning. In *WSDM* (pp. 223–232).
- Evangelatos, O., Samarasinghe, K., & Rolim, J. (2013). Syn desi: A framework for creating personalized smart environments using wireless sensor networks. In *2013 IEEE international conference on distributed computing in sensor systems* (p. 325-330).
- Freeman, A., Adams Becker, S., Cummins, M., McKelroy, E., Giesinger, C., & Yuhnke, B. (2016). *The NMC horizon report: 2016: Museum edition*. The New Media Consortium, Austin, Texas.
- Friess, P. (2013). *Internet of things: converging technologies for smart environments and integrated ecosystems*. River Publishers.
- Gao, H., Tang, J., Hu, X., & Liu, H. (2013). Exploring temporal effects for location recommendation on location-based social networks. In *Proceedings of the 7th ACM conference on recommender systems* (pp. 93–100). ACM.
- Gribaudo, M., Iacono, M., & Levis, A. H. (2017). An IoT-based monitoring approach for cultural heritage sites: The Matera case. *Concurrency and Computation: Practice and Experience*, 29(11).
- Grieser, K., Baldwin, T., & Bird, S. (2007). Dynamic path prediction and recommendation in a museum environment. In *Proceedings of the workshop on language technology for cultural heritage data (LaT-eCH 2007)* (pp. 49–56).
- Griesner, J.-B., Abdessalem, T., & Naacke, H. (2015). POI recommendation: Towards fused matrix factorization with geographical and temporal influences. In *Proceedings of the 9th ACM conference on recommender systems* (pp. 301–304). ACM.
- Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, 29(7), 1645 - 1660.
- Guy, I. (2015). The role of user location in personalized search and recommendation. In *Proceedings of the 9th ACM conference on recommender systems* (pp. 236–236). ACM.
- Hashemi, S. H., Clarke, C. L. A., Kamps, J., Kiseleva, J., & Voorhees, E. M. (2016). Overview of the trec 2016 contextual suggestion track. In *Proceeding of text retrieval conference*

- (TREC).
- Hashemi, S. H., Hupperetz, W., Kamps, J., & van der Vaart, M. (2016). Effects of position and time bias on understanding onsite users' behavior. In *Proceedings of the 2016 ACM on conference on human information interaction and retrieval* (pp. 277–280). ACM.
- Hashemi, S. H., & Kamps, J. (2017a). Skip or stay: Users' behavior in dealing with onsite information interaction crowd-bias. In *Proceedings of the 2017 ACM on conference on human information interaction and retrieval*. ACM.
- Hashemi, S. H., & Kamps, J. (2017b). Where to go next?: Exploiting behavioral user models in smart environments. In *Proceedings of the 25th conference on user modeling, adaptation and personalization* (pp. 50–58). ACM.
- Hashemi, S. H., Williams, K., El Kholy, A., Zitouni, I., & Crook, P. (2018a). Impact of domain and user's learning phase on task and session identification in smart speaker intelligent assistants. In *Proceedings of the 27th acm international conference on information & knowledge management*.
- Hashemi, S. H., Williams, K., El Kholy, A., Zitouni, I., & Crook, P. (2018b). Measuring user satisfaction on smart speaker intelligent assistants using intent sensitive query embeddings. In *Proceedings of the 27th acm international conference on information & knowledge management*.
- Herlocker, J. L., Konstan, J. A., Terveen, L. G., & Riedl, J. T. (2004). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems*, 22(1), 5–53.
- Hernández-Muñoz, J. M., Vercher, J. B., Muñoz, L., Galache, J. A., Presser, M., Gómez, L. A. H., & Pettersson, J. (2011). Smart cities at the forefront of the future internet. In *The future internet assembly* (pp. 447–462).
- Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2012). Improving neural networks by preventing co-adaptation of feature detectors. *CoRR*, abs/1207.0580.
- Johnson, L., Adams Becker, S., Estrada, V., & Freeman, A. (2015). *The NMC horizon report: 2015 museum edition*. The New Media Consortium, Austin, Texas.
- Karatzoglou, A., Amatriain, X., Baltrunas, L., & Oliver, N. (2010). Multiverse recommendation: N-dimensional tensor factorization for context-aware collaborative filtering. In *Proceedings of the fourth ACM conference on recommender systems* (pp. 79–86). ACM.
- Lanir, J., Kuflik, T., Dim, E., Wecker, A. J., & Stock, O. (2013). The influence of a location-aware mobile guide on museum visitors' behavior. *Interacting with Computers*, 25(6), 443–460.
- Liu, X., & Aberer, K. (2013). SoCo: A social network aided context-aware recommender system. In *Proceedings of the 22nd international conference on world wide web* (pp. 781–802). ACM.
- Lucchese, C., Perego, R., Silvestri, F., Vahabi, H., & Venturini, R. (2012). How random walks can help tourism. In *Proceedings of advances in information retrieval: 34th european conference on IR research, ECIR 2012*.
- Mighali, V., Del Fiore, G., Patrono, L., Mainetti, L., Alletto, S., Serra, G., & Cucchiara, R. (2015). Innovative IoT-aware services for a smart museum. In *Proceedings of the 24th international conference on world wide web* (pp. 547–550). ACM.
- Park, M.-H., Hong, J.-H., & Cho, S.-B. (2007). Location-based recommendation system using bayesian user's preference model in mobile devices. In *International conference on ubiquitous intelligence and computing* (pp. 1130–1139).
- Perera, C., Zaslavsky, A., Christen, P., & Georgakopoulos, D. (2014). Context aware computing for the internet of things: A survey. *IEEE Communications Surveys & Tutorials*, 16(1), 414–454.
- Rao, A. S., Sharma, A. V., & Narayan, C. S. (2017). A context aware system for an IoT-based smart museum. In *2017 2nd international multidisciplinary conference on computer and energy science (SpliTech)* (p. 1-5).
- Rendle, S., Gantner, Z., Freudenthaler, C., & Schmidt-Thieme, L. (2011). Fast context-aware recommendations with factorization machines. In *Proceedings of the 34th international*

- ACM SIGIR conference on research and development in information retrieval* (pp. 635–644). ACM.
- Scholz, C., Illig, J., Atzmueller, M., & Stumme, G. (2014). On the predictability of talk attendance at academic conferences. In *Proceedings of the 25th ACM conference on hypertext and social media* (pp. 279–284). ACM.
- Shokouhi, M., Ozertem, U., & Craswell, N. (2016). Did you say U2 or YouTube?: Inferring implicit transcripts from voice search logs. In *WWW* (pp. 1215–1224).
- Sornalatha, K., & Kavitha, V. R. (2017). IoT based smart museum using bluetooth low energy. In *2017 third international conference on advances in electrical, electronics, information, communication and bio-informatics (AEEICB)* (p. 520-523).
- Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(1), 1929–1958.
- Strohbach, M., Ziekow, H., Gazis, V., & Akiva, N. (2015). Towards a big data analytics framework for IoT and smart city applications. In F. Khafa, L. Barolli, A. Barolli, & P. Papajorgji (Eds.), (pp. 257–282).
- Wang, G., Zhang, X., Tang, S., Zheng, H., & Zhao, B. Y. (2016). Unsupervised clickstream clustering for user behavior analysis. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 225–236). ACM.
- Wecker, A. J., Lanir, J., Kuflik, T., & Stock, O. (2011). Pathlight: Supporting navigation of small groups in the museum context. In *Proceedings of the 13th international conference on human computer interaction with mobile devices and services* (pp. 569–574). ACM.
- Xiao, X., Zheng, Y., Luo, Q., & Xie, X. (2010). Finding similar users using category-based location history. In *Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems* (pp. 442–445). ACM.
- Ye, M., Yin, P., Lee, W.-C., & Lee, D.-L. (2011). Exploiting geographical influence for collaborative point-of-interest recommendation. In *Proceedings of the 34th international ACM SIGIR conference on research and development in information retrieval* (pp. 325–334). ACM.
- Ying, J. J.-C., Lu, E. H.-C., Kuo, W.-N., & Tseng, V. S. (2012). Urban point-of-interest recommendation by mining user check-in behaviors. In *Proceedings of the ACM SIGKDD international workshop on urban computing* (pp. 63–70). ACM.
- Zanella, A., Bui, N., Castellani, A., Vangelista, L., & Zorzi, M. (2014). Internet of things for smart cities. *IEEE Internet of Things Journal*, 1(1), 22-32.
- Zheng, V. W., Cao, B., Zheng, Y., Xie, X., & Yang, Q. (2010). Collaborative filtering meets mobile recommendation: A user-centered approach. In *Proceedings of the twenty-fourth AAAI conference on artificial intelligence* (pp. 236–241). AAAI Press.
- Zheng, V. W., Zheng, Y., Xie, X., & Yang, Q. (2010). Collaborative location and activity recommendations with GPS history data. In *Proceedings of the 19th international conference on world wide web* (pp. 1029–1038). ACM.
- Zheng, Y., Zhang, L., Xie, X., & Ma, W.-Y. (2009). Mining interesting locations and travel sequences from GPS trajectories. In *Proceedings of the 18th international conference on world wide web* (pp. 791–800). ACM.
- Zhuang, J., Mei, T., Hoi, S. C., Xu, Y.-Q., & Li, S. (2011). When recommendation meets mobile: Contextual and personalized recommendation on the go. In *Proceedings of the 13th international conference on ubiquitous computing* (pp. 153–162). ACM.
- Zoeter, O. (2015). Recommendations in travel. In *Proceedings of the 9th ACM conference on recommender systems* (pp. 234–234). ACM.