Inverse Foraging: Inferring Users' Interest in Pervasive Displays

MARIA L. MONTOYA FREIRE, Aalto University, Finland
ANTTI OULASVIRTA, Aalto University, Finland
MARIO DI FRANCESCO, Aalto University, Finland

Users' engagement with pervasive displays has been extensively studied, however, determining how their content is interesting remains an open problem. Tracking of body postures and gaze has been explored as an indication of attention; still, existing works have not been able to estimate the interest of passers-by from readily available data, such as the display viewing time. This article presents a simple yet accurate method of estimating users' interest in multiple content items shown at the same time on displays. The proposed approach builds on the information foraging theory, which assumes that users optimally decide on the content they consume. Through *inverse* foraging, the parameters of a foraging model are fitted to the values of viewing times observed in practice, to yield estimates of user interest. Different foraging models are evaluated by using synthetic data and with a controlled user study. The results demonstrate that inverse foraging accurately estimates interest, achieving an R^2 above 70% in comparison to self-reported interest. As a consequence, the proposed solution allows to dynamically adapt the content shown on pervasive displays, based on viewing data that can be easily obtained in field deployments.

CCS Concepts: • **Human-centered computing → User models**; *Ubiquitous and mobile computing*.

Additional Key Words and Phrases: parameter estimation, inverse modeling, information foraging theory, pervasive displays

ACM Reference Format:
Maria L. Montoya Freire, Antti Oulasvirta, and Mario Di Francesco. 2021. Inverse Foraging: Inferring Users' Interest in Pervasive Displays. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 3, Article 122 (September 2021), 18 pages. https://doi.org/10.1145/3478103

1 INTRODUCTION

Pervasive displays are increasingly employed to convey diverse information in different settings, from private spaces to urban environments [54]. Researchers have investigated how to draw users' attention toward the displays, primarily by creating new interaction methods: multi-touch displays [19, 46] and the use of external devices such as cell phones [9, 22, 27, 45] or eye-trackers [26, 61, 64], to name a few. This article addresses a hard problem in the design and deployment of pervasive displays: how to know what content users are interested in [4]. The key idea is that users interested in certain content will devote more time to peruse it. Instead, users will attend less to displays that show content they do not consider interesting [37, 41]. However, when *multiple* content items are presented, how can one estimate users' interest without asking them or tracking their gaze?

Although extensive research has been conducted to characterize users' engagement with pervasive displays, less effort has been directed towards methods suitable for the non-interactive scenarios that are prevalent in many applications [6]. One solution is to measure users' *attention* with tracking devices together with computer-vision methods. The main idea behind this approach is to determine users' level of attention on the basis of their head

Authors' addresses: Maria L. Montoya Freire, Aalto University, Espoo, Finland, maria.montoyafreire@aalto.fi; Antti Oulasvirta, Aalto University, Espoo, Finland, antti.oulasvirta@aalto.fi; Mario Di Francesco, Aalto University, Espoo, Finland, mario.di.francesco@aalto.fi.

2474-9567/2021/9-ART122
https://doi.org/10.1145/3478103

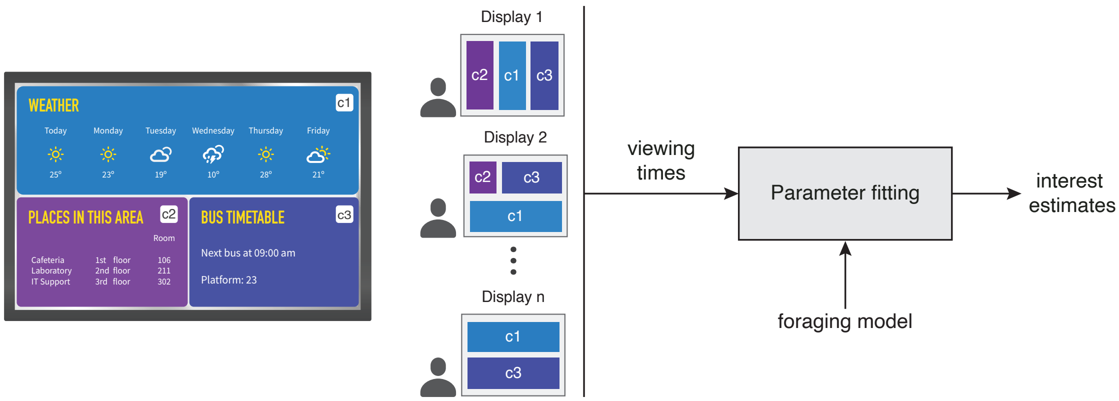


Fig. 1. Overview of the proposed approach. Inverse foraging refers to the process of inferring users' interest in content from the corresponding viewing data. This method takes as input both viewing times and information on what was shown on different displays. Users' interest is estimated by fitting the interest parameters of a foraging model to collected data.

posture and gaze direction with respect to the screen [2, 36]. Precisely, the use of tracking devices leverage mobility (i.e., of body parts or eye gaze) to derive attention. However, attention itself is an indication of noticing something (e.g., a flashing light) and is not immediately connected to interest, particularly, in visual content.

User modeling is a widely used approach in human-computer interaction (HCI) to inform interface design and optimize systems to improve performance given certain inputs [14]. It has been as well applied to generate more engaging display content through models that link users' attention to their interest [39]. Specifically, so-called *foraging models* have been employed to characterize the behavior of a user deciding on which content to direct attention to. These models, rooted in economics and biology, are based on a principle of optimality: from among a set of sources, the user selects the one that provides the greatest benefit – for instance, the source that yields the most informative value. The rationale of these models is that users' preferences and goals affect their decision to attend to a particular content item, as does the time available. Foraging models highlight that visual attention is not related to interest in a straightforward manner. In fact, the time spent attending to content is affected by several other factors, such as the amount of information contained in it and its visual prominence.

Furthermore, a common approach to solve problems in HCI has been to build forward models [44]. In *forward modeling*, human-like data is generated from a model. To produce realistic data, the corresponding parameters must be assigned meaningful values, which often entails employing findings from previous studies or manually setting the values. In some cases, values assigned in this way might not suffice for the study at hand, thereby affecting model accuracy. To overcome this issue, researchers have also employed *inverse modeling*, in which the optimal values for model parameters are found by fitting the model to observed data [24]. In other words, forward modeling involves building a model to obtain data, whereas inverse modeling employs data to infer parameters for a model [58]. Recent studies [15, 21, 23, 24] have demonstrated that complex interaction data can be handled by modern inverse modeling methods.

Motivated by these considerations, this article addresses users' engagement with displays by adopting an inverse modeling approach. To the best of the authors' knowledge, this approach has not been applied before, particularly, in the context of pervasive displays (Section 2). Specifically, this work proposes *inverse foraging*, a method to estimate user interest in the individual content items shown on a display. To this end, it employs foraging models for fitting data on viewing times (see Figure 1). As a result, the proposed approach can precisely determine the content considered interesting for users among all the different items shown on the display. This is a simple yet accurate method of inferring interest, in that it does not require any user interaction but only relies on aggregate viewing times. Accordingly, inverse foraging can be meaningfully applied to the challenging case of non-interactive displays, whose content is not explicitly requested by users. The proposed approach allows to dynamically adapt the content shown on pervasive displays based on viewing data, as a form of implicit feedback on the corresponding interest. For this reason, it could be easily adopted and deployed in real settings.

This work establishes the following contributions.

- Two foraging models are proposed to characterize the behavior of users looking at a pervasive display (Section 3). These models are expressive, as they capture the distinctive features of tiled display layouts. They are also simple, thus, suitable for inversion.
- A method to estimate users' interest is devised by considering the viewing process of a display as an inverse problem (Section 4). In particular, the proposed method takes viewing data as input and derives the values of the parameters in the foraging models to maximize prediction accuracy.
- The quality of the model is examined through simulation with synthetic data (Section 5). The evaluation shows that an accurate characterization is achieved in practice with only a few layouts.
- A user study is carried out to validate the model with real-world data (Section 6). The results demonstrate that the proposed method can effectively estimate users' interest.
- Finally, guidelines for applying inverse foraging in real settings are provided (Section 7). They also include a practical method to obtain viewing data from display audience with limited instrumentation, without relying on devices carried by users.

2 RELATED WORK

Inverse modeling is increasingly applied in HCI research to better exploit predictive models in explaining and improving interaction. The discussion below outlines several approaches proposed in the literature that are relevant to this work.

2.1 Inverse Models in HCI Research

Prior work has applied inverse models to characterize users' actions in several contexts, such as keyboard layouts [21] and task interleaving [15]. In addition, research has been carried out to explore new approaches to estimate models' parameters from data. For instance, Approximate Bayesian Computation (ABC) [24] has been proposed to compute posterior distributions for parameters, which can aid in understanding the identifiability of models. When recently compared to non-Bayesian methods of inverse modeling in the context of cognitive models, the ABC approach demonstrated itself to be efficient [23]. While prior work attests to the efficiency of inverse models in multiple contexts, scenarios with pervasive displays have not been considered. Hence, this article elaborates on inverse models to address the challenging case in which multiple content items are shown on a display.

2.2 Information Foraging

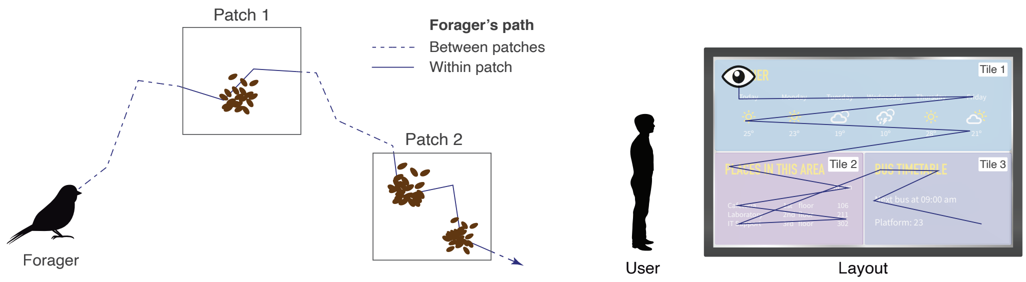
The information foraging theory has been applied to understand interactive behavior that can be described as choices. The assumption is that users (foragers) search sources of information and choose one with the goal of maximizing information gain, in a manner similar to how animals search for food [47]. Emerging evidence suggests that humans adopt foraging behavior not only to search information in external environments but also inside the mind [17, 59]. Previous studies have explored the use of foraging models in various domains, including image retrieval [20, 33], recommender systems [53], Web search [11, 48, 60], and programming tasks [25, 28, 30, 31]. Montoya et al. [39] addressed the problem of engaging users with pervasive displays by applying the information foraging theory. They introduced the *display foraging model*, which enables maximizing users' information gain during reading of content shown on a display with tiled layouts. The model is built under the assumption that users, having limited time and attention, optimize their selection of content. While the proposed approach here is mainly based on such information foraging models, the main idea is to estimate parameters from data; thereby, one can obtain a better estimate of the user's interest with regard to the individual content items displayed on a screen.

2.3 Interest Models

There is a growing body of literature on the use of click data to estimate user interest in the context of Web search [5, 34, 56, 65]. Utilizing user's behavioral data has been proven effective to improve the Web search results [1]. Shen et al. [56] introduced a framework in which personalized Web search results are provided by means of a collaborative-filtering approach, which involves collecting information from many users about their behavior and item preferences, to serve as seeds for new recommendations. Also employing a collaborative-filtering approach, Liu et al. [34] combined this technique with data obtained from user profiles to generate personalized news recommendations. Instead, Qiu and Cho [51] developed an algorithm to learn users' interests from their click history by considering their topic preferences. While the findings from using click-behavior data are promising, clicks do not necessarily reflect interest in a given item; for instance, the user might click on an item because of its title but then find the content unsuitable and not interesting, which leads the user to express a negative preference for it. As a consequence, other studies [35, 63] have incorporated alternative features into the solution, such as dwell time, which offers insight into how long the user reads the content. To address such issues in the context of pervasive displays, where clicks are not available, the study reported upon here took a relatively simple approach, which requires only collecting users' viewing time data for estimation of their interest with respect to a set of content items. Moreover, the proposed approach is practical as it does not require any user input.

2.4 Measuring Attention to Pervasive Displays

Several studies have investigated techniques to measure user attention directed toward a pervasive display. For instance, tracking devices (e.g., Kinect) have been used by researchers to analyze user behavior by detecting facial features [13, 42]. Thereby, it is possible to determine whether users are looking at a display on the basis of their head orientation or the direction of their gaze. Schiavo et al. [52] developed an approach to estimate users' interest from a set of conditions (i.e., distance from the screen, head orientation, and social context). Certainly, this approach allows to estimate user attention to a screen; however, it only considers information connected with user behavior, while no consideration is given to the content shown. In other works, researchers have achieved promising results with solutions that replace tracking devices with off-the-shelf cameras to estimate the user's attention. For instance, Asteriadis et al. [3] presented a system that can estimate the level of frustration or attention by combining head posture with gaze features. Moreover, AggreGaze [57] employed an appearance-based gaze-estimation method specifically to estimate user attention to a display spatio-temporally. This approach indeed provides more accurate information on what part of the display the user is focusing on, but its use has been limited to an evaluation only involving video content. In specialist domains, systems have been developed to estimate user attention in real-world contexts. For instance, MyAds [10] has employed RFID tags to collect data and, thereby, produce display-relevant advertisements in line with user profiles. Similarly, BlueScreen [55] exploits the presence of mobile devices with enabled Bluetooth functionality to advertise content judged interesting to users. In addition, a multifaceted approach proposed by Müller and Krüger [40] employs several features for estimation of user's interest, such as time, location, and the number of people looking at a given display. All these solutions require user intervention, in that one must be carrying one's phone or RFID-tagged items, which might not be practical in some scenarios. These and the other previous studies have relied on external devices to measure interest and adapt the content shown on that basis. With particular devices, the information



(a) (b)

Fig. 2. Illustration of the patch model in the IFT for (a) a food foraging scenario and (b) its application to tiled display layouts. A bird forages areas containing food (namely, patches) in the environment, spending time for both finding and consuming them. Likewise, a user peruses tiles in a display layout, by looking at the different content items therein.

obtained is limited to physical features of a user's response to the display, hence the system does not provide any insight related to the content itself. Moreover, a recent study [38] presented a novel approach that can provide viewer-centric digital signage analytics (e.g., number of viewers per content item and time spent on a display). The approach consists of combining traditional sign data analytics with user mobility simulations. While the findings demonstrate the potential of this solution, it could only provide insights to improve content scheduling and how to effectively place displays. This article, in contrast, mainly focuses on the content shown and on an associated numerical value for purposes of distinguishing between content items that are interesting, according to users, and those that are not.

3 FORAGING MODELS FOR PERVASIVE DISPLAYS

This section first introduces the key concepts behind the information foraging theory that is employed to model how people process and consume informative content. It then presents two different models that apply the related theory to the specific context of pervasive displays. In the considered scenario [39], the area of these displays is divided into separate *tiles*, showing different content items and organized into a space-filling arrangement (i.e., a *layout*).

3.1 Information Foraging Theory

The Information Foraging Theory (IFT) [47] explains the behavior of an agent (i.e., a user) seeking information from a set of sources by drawing upon concepts and techniques originally developed in behavioral ecology. The analogy is based on how organisms forage for food in an environment subject to certain constraints. The key idea is that these organisms seek an *optimal* foraging strategy: they aim at maximizing their energy intake given the availability of food and the effort in obtaining it. Likewise, users spend a certain time in consuming information sources to maximize the efficiency in accomplishing a task. This is explained based on the assumption that users are *rational*, willingly selecting actions to achieve a goal, based on their knowledge [43].

There are different types of foraging models. The *patch* model is the one most relevant to this work: the environment is divided into different areas (indeed called patches) containing food items, each offering a certain reward (Figure 2a). Patches can vary in terms of prevalence (their relative occurrence) and profitability (the amount of food / reward). A forager looks for a patch and stays therein to consume its content, then continues to another patch and so on. The optimization problem here is finding the optimal amount of time to spend *within* a patch before going forward (i.e., moving *between* patches). This can be formulated in terms of the so-called *gain*, a measure of the value resulting from foraging. Individual patches have their own gain function, which is generally characterized by diminishing returns: the rate of gain reduces over time and eventually remains constant. For this reason, logarithmic gain functions in the form $g(t) = a \cdot \ln(t) - b$ are widely used, wherein a is a scaling factor and b the offset of the function on the vertical axis, respectively. Diminishing returns are due to the fact that rewards are limited and not replenishable. For instance, a bird foraging berries in a bush can eventually eat all of them; while the gain increases quickly at the beginning as there are plenty of ripe berries that are easy to reach, it later decreases as fewer, less ripe berries remain in locations that are more difficult to get to. Accordingly, the forager should remain in a patch only as long as the marginal value of the related gain function is higher than the average rate of gain of the environment.

3.2 The Display Forager

The *display* forager applies the IFT to the case of tiled display layouts to characterize non-interactive, multi-content pervasive displays, similar to [39]. In these settings, the layout is considered as the environment, tiles are patches, and different content items are types of rewards (Figure 2b). More formally, a given layout l is composed by K tiles, individually denoted with an index k, with $1 \leq k \leq K$. Each tile is characterized by different parameters: the coefficients a_k and b_k of a logarithmic gain function, which are assumed to be the same for all tiles in a layout (i.e., $a_k = a_l$ and $b_k = b_l, \forall\, k : 1 \leq k \leq K$); an interest value i_k; the geometry-related parameters s_k for the area and r_k for the aspect ratio of tile k, respectively. Individual content items can be shown on tiles with different geometry, in which case they are adapted according to the tile size and aspect ratio.

The display forager derives the total time a user spends by looking at a layout as a function of the tiles contained therein. Specifically, the time t_k spent by a certain user to inspect tile k is defined as follows [39, Section 2.3.2]:

$$t_k = \frac{a_l}{i_k} + \frac{b_l \cdot s_k}{r_k} \tag{1}$$

The right hand-side of the equation is composed of two terms. The first relates to the scaling factor of the gain function, which is affected by the interest as the profitability in the IFT, thereby explaining the inversely proportional relationship [47, p. 34]. The second term refers to the offset of the gain function and the geometry of the tile: the area describes its visual prominence, while the aspect ratio is an indication of how it is cumbersome to inspect content that deviates from a horizontal arrangement [12].

The total time T_l of perusing a layout can then be calculated as the sum of the values t_k for the individual tiles shown therein:

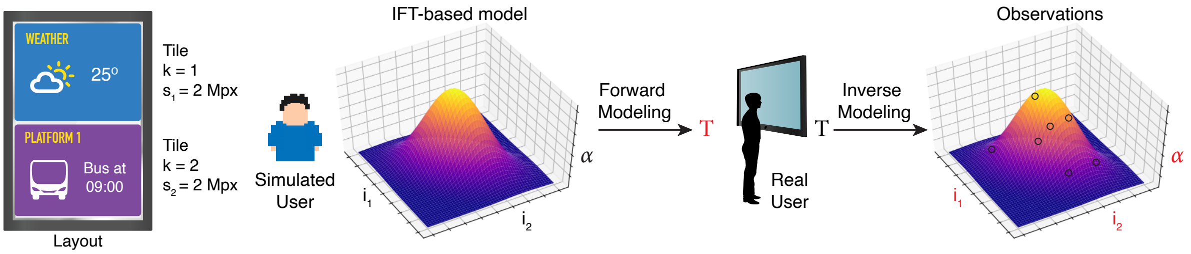
$$T_l = \sum_{k=1}^{K} \left(\frac{a_l}{i_k} + \frac{b_l \cdot s_k}{r_k} \right) \tag{2}$$

With reference to the IFT, this choice considers the overall foraging time as the sum of the time within the patches in the environment, under the assumption that the time between patches is negligible (since the tiles are adjacent to each other).

3.3 The Tile Forager

The *tile* forager is a simplified version of the display forager, which does not rely on the parameters associated with the information gain (i.e., a_l and b_l) to derive the total viewing time.

In detail, the tile forager assumes that the time spent on tile k jointly depends on the area s_k and the interest value i_k associated with a content item shown therein. Such a relation indicates that the prominence of the tile affects the viewing time proportionally to the interest of the user in a particular content item. The total time T_l for perusing layout l is then calculated as a linear combination of the time for an individual tile, weighted by a



(a) (b) (c)

Fig. 3. Example of the tile forager as a forward and an inverse model. (a) Simple layout used for illustration purposes, composed of two tiles of the same size; the information about the layout (i.e., the content associated to the tiles and their geometry) are always known in the considered context. (b) In *forward* modeling, all the other parameters (i.e., the alpha coefficient and the interest for individual content items) are also assumed to be known and the total viewing time is derived through Eq. (3). (c) In *inverse* modeling, only the total viewing time is known (through observations) and black-box optimization is applied to derive the unknown parameter values.

coefficient α_l:

$$T_l = \alpha_l \sum_{k=1}^{K} s_k \cdot i_k \tag{3}$$

As a normalization factor that allows more flexibility in the possible choices of interest values across tiles.

The tile forager assumes that the size of the tile is the primary factor: the larger the tile, the more informative content and the greater the attention. This model further assumes that the relationship between size and interest is directly proportional, unlike that in the display forager. Such relaxations depart from the theoretical foundation of the IFT [47]; however, they are beneficial in that the model is very simple, with only a few parameters. This makes inversion easier and prevents overfitting. As will be shown in Section 5, accurate predictions are obtained notwithstanding the low number of parameters.

4 INVERSE MODELING

The previous section discussed information foraging in the context of *forward* modeling. Accordingly, forward modeling builds a mathematical abstraction that generates human-like data [44]. For instance, the IFT-based models introduced in the previous section can be applied to predict the total viewing time by knowing all the parameters describing a layout: the geometry of tiles, their content items and the related interest (Figures 3a and 3b). The viewing time, in turn, can be leveraged for the computational design of informative display layouts that maximize the information gain [39]. Such an approach is solid, as grounded on theories that accurately describe human behavior in realistic settings by drawing from cognitive science. However, it is not at all straightforward to set the values of the parameters for specific settings, as this may require conducting preliminary studies that are cumbersome to carry out or difficult to generalize.

A different option is given by *inverse* modeling. Broadly speaking, inverse modeling is a data-driven approach to customize a mathematical representation of a system for a certain scenario. More specifically, inverse modeling applies computational parameter-fitting methods to estimate the actual values in a target model that best explain the user behavior expressed by the data [58]. In contrast with forward modeling, here parameter values are unknown and are treated as variables, which are then instantiated on the basis of available observations – realizations of the model output through a quantitative characterization of user behavior (Figure 3c). Inverse modeling is also practical, as long as the output of the model can be measured with some accuracy in an unobtrusive way. As for the above-mentioned IFT-based models, inversion only requires obtaining the time spent by a user in perusing a certain layout. This can be achieved, for instance, with the method described in Section 7.

A preliminary step for model inversion is data collection based on observations of actual user behavior. In the considered context, an observation consists of measuring the user-specific viewing time for a given layout. A dataset can be built by considering multiple observations: as distinct displays showing different layouts that do not change over time, a single display that cycles through different layouts, or both. The exact realization does not matter, as long as there are enough layouts and variety in their configurations (i.e., diversity in the geometry and content of tiles). Otherwise, the model parameters might not be correctly recovered, as it will be further explained in the next section.

Once viewing times are available, it remains to apply parameter fitting. The actual parameters to be estimated depend on the considered model. The interest i_k for individual content items is derived for both IFT-based models; furthermore, the a_l and b_l coefficients associated with the gain function are obtained for the *display* forager, the α coefficient for the *tile* forager. Moreover, different methods can be used for model fitting, including those for black-box optimization that do not even require knowing the analytical expression of a function – e.g., grid search or Bayesian optimization [24]. Such an optimization consists in minimizing the difference between model's prediction and the observed data. Widely used methods include: Powell [49], employing two search vectors to define the direction for searching parameter values; Nelder-Mead [29], leveraging a modified version of the simplex method and iterative reshaping of the search space; and Differential Evolution (DE) [50], applying a genetic algorithm. A preliminary study revealed that both Powell and Nelder-Mead were not effective in the considered scenario, whereas DE achieved considerably smaller prediction errors when its hyperparameters were selected through Bayesian optimization. As a consequence, the latter was selected to perform parameter fitting.

5 ASSESSING PARAMETER RECOVERY

Before fitting the models to empirical data, it is essential to check that model parameters can actually be estimated from such data. This step is called *parameter recovery* [16]. To this end, synthetic data were first generated and fed to the models, then an evaluation was carried out to assess whether the parameters generated can be inferred from said data [62]. Accordingly, the model that generated the data is fitted, and effectiveness is quantified in terms of accuracy. Ideally, there should be a strong correlation between the true values and the recovered parameters. Indeed, parameter recovery is an important (though often overlooked) step in modeling, in that it allows to discover possible issues affecting model quality. The following subsections describe the method to generate the synthetic data and present the results obtained for the considered cases.

5.1 Data Generation

Diverse realistic simulated scenarios were created for evaluation wherein different numbers of layouts were viewed by different users. Synthetic data were created by considering various numbers of layouts and also several types of content items. In total, four datasets were created, each containing several layouts with distinct number of tiles (between 2 and 6). As for the type of content items, 8 options were selected in line with the assumption that the content would be shown in a university setting. The content was randomly assigned to individual tiles for all the layouts. Also, the body of data was created under the assumption that there are several users standing in front of the display and that some content items are found to be more interesting than others. Accordingly, the data included the time devoted to each layout, with its value set within the range of 1.9 to 6.6 seconds, which is consistent with the values reported in prior work [18, 39]. The interest ratings were employed as the ground truth, with values defined in a range of 1 to 7, for content items deemed "least interesting" and "very interesting", respectively. The foraging models were assessed for estimating interest under two evaluation scenarios: five content items with 6 and 14 layouts as well as eight content items with 14 and 20 layouts.

Table 1. Evaluation results obtained for the individual models with synthetic data.

		Tile forager				**Display forager**			
No. of layouts	Content items	AIC	BIC	RMSE	R^2	AIC	BIC	RMSE	R^2
6	5	27.27	20.64	1.342	0.789	36.59	27.44	1.008	0.771
14	5	41.39	30.72	1.015	0.816	72.57	56.32	1.378	0.826
14	8	60.70	51.41	1.425	0.832	97.37	88.70	2.376	0.850
20	8	70.79	61.90	1.258	0.848	115.98	107.94	1.662	0.872
	Mean	50.04	41.17	1.260	0.821	80.63	70.10	1.606	0.830
	SD	19.47	18.84	0.177	0.025	34.32	35.53	0.579	0.043

5.2 Implementation

The IFT-based models were implemented in Python. Model inversion was realized through the DE algorithm as available in the `scipy.optimize` library. The hyperparameters of DE were fitted through the tools for Bayesian optimization in the `scikit-optimize` library[1]. The code was optimized for parallel execution, and model-fitting was carried out on a high-performance computing cluster.

5.3 Results

The following reports accuracy in terms of the following metrics: root mean square error (RMSE), expressed in the same units as the interest ratings (i.e., on the 1–7 scale); and R^2, which explains how well the model fits the data, from 0% to 100%. In addition, the following metrics are calculated to compare the models: the Akaike Information Criterion (AIC) as an estimate of how well the model fits the data without overfitting; and the Bayesian Information Criterion (BIC), used for model selection.

Before reporting the overall findings, the results obtained for the considered scenarios are described below. The first features five content items arranged in up to 14 distinct layouts. Table 1 shows examples wherein the two foraging models were fitted to synthetic user data from six and 14 layouts. When using six layouts, the models reach similar levels of accuracy, with R^2 above 0.75. However, the prediction error is smaller for the display forager, with an RMSE value of 1.008. The R^2 increases when 14 layouts are used. Here, the tile forager obtains a value of 0.816 and the RMSE decreases to 1.015. In contrast, the display forager obtains a higher R^2 value, 0.826, while the RMSE also increases, by 0.363, from that with the tile forager. From these results, it can be concluded that having more displays to draw data from increases the accuracy of the inverse modeling.

Moreover, the number of content items per layout was varied. When more content items are presented, the display forager achieves higher accuracy than the tile forager, with $R^2 = 0.872$ as shown in Table 1. However, the tile forager obtains lower RMSE values in both cases, and for 20 layouts that model actually performs better, with a minimum prediction error of 1.258 units. In contrast, the display forager shows a prediction-error value of 2.376 when 14 layouts are used.

Both models were effective for inversion, as shown by the aggregated results. The tile forager performed slightly better, as it obtained lower values of both prediction error (an average RMSE of 1.260) and AIC (50.04 on average), indicating less overfitting than the display forager. A similar pattern can be observed in the BIC for the two models. Figure 4 shows sample results obtained by the tile forager which obtained lower prediction errors on average for the two considered scenarios.

[1] `https://scikit-optimize.github.io/dev/index.html`

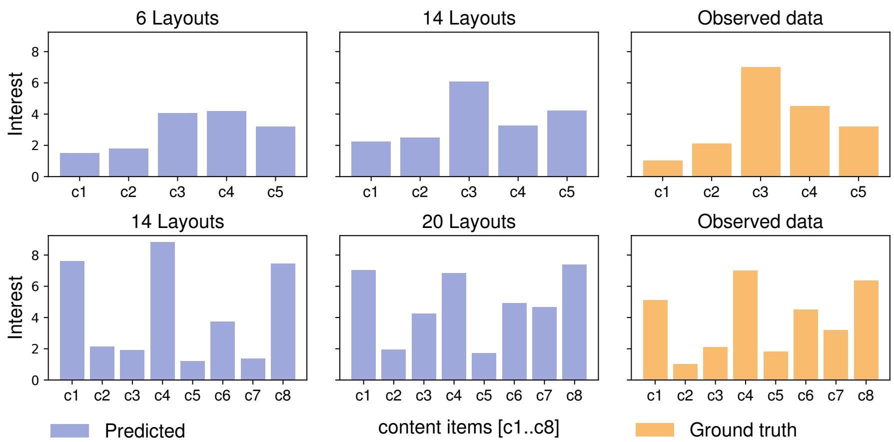


Fig. 4. Parameter recovery, with selected results from inverting data synthetically generated by the tile forager. The first row (top) shows the predictions obtained by the model for five content items. The second row (bottom) depicts the model's predictions for eight content items. The results show that the model can estimate the original interest parameters with very satisfactory accuracy.

6 EMPIRICAL EVALUATION

A controlled user study was conducted to evaluate the accuracy of inverse foraging with real viewing data on tiled display layouts. To this end, a free-viewing task [39] was employed wherein participants were asked to view different layouts with multiple content items, one by one, pressing a button when done with each. The benefit of organizing the research as a controlled study was to mitigate the influence of confounding variables and to gain higher precision in (viewing time) measurements than field studies would permit [7]. After viewing the layouts, the participants were asked to fill in a rating form to express their interest in the content presented. Accordingly, the interest ratings served as the ground truth to assess the accuracy of the inverse foraging.

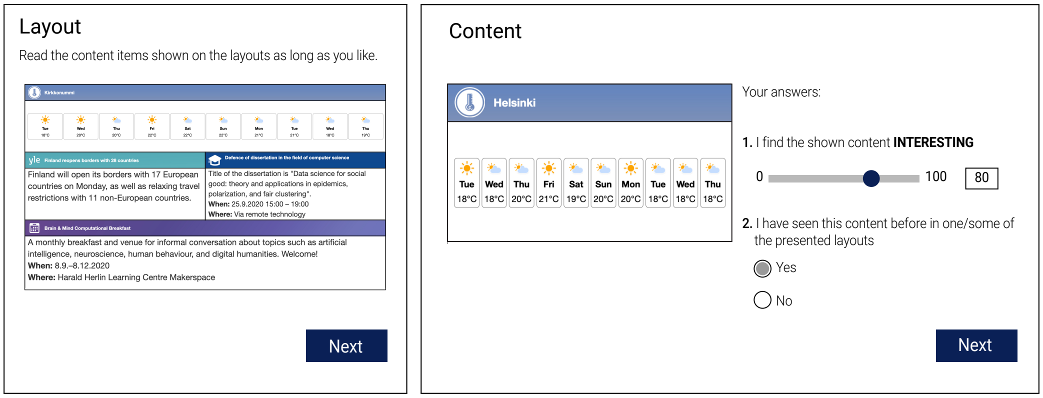
6.1 Participants

Sixteen participants were recruited at the authors' institution (10 male and 6 female). The participants had diverse educational background and their ages ranged between 25 and 33 years (M = 28.43, SD = 2.70). Recruitment was performed via several messaging platforms; participation in the study was voluntary and subject to informed consent.

6.2 Materials

For the study, a Web-based application was implemented to generate layouts, similar to the one described by Montoya et al. [39]. The application randomly generated 200 distinct layouts, with each layout containing between two and six tiles. Those layouts were then exported as high-resolution images for use in the survey. The study was conducted in a university context, therefore content items were created from data sources related to the university. To this end, several types of content were considered, including upcoming events, local news, and posts on social media. A probability distribution was the criterion employed to assign content items to individual tiles.

A Web-based survey was realized by following a within-subject design, with two variables considered for analysis: the viewing time of a layout and the user's interest in certain content items. The first variable was measured as the duration of viewing a layout, the second was expressed in terms of participant's ratings. The



(a) (b)

Fig. 5. Screenshots of the interface employed in the user study, consisting of two phases. (a) In the first one, users were shown layouts with multiple content items and asked to press a button once they found them no longer interesting. (b) In the second phase, ground-truth measurements for interest were obtained in a rating task.

survey showed a sequence of five layouts and asked 11 questions. The interface was easy to use: participants merely had to press a button to show the next layout or question (Figure 5). The study was carried out online, therefore participants used their own equipment as long as it was either a laptop or a desktop (workstation) with a large-enough widescreen display.

6.3 Procedure

Participants were informed about the data collected during the study and were given a URL to access the survey. The participant had to read the instructions to use the tool and was asked for informed consent before starting the experiment. There was no time limit to complete the study, hence participants could perform the tasks at their own pace.

The participants had to perform two tasks to complete the study. Accordingly, the study was divided into two phases: self-paced viewing and content ratings.

6.3.1 Phase I: Self-paced viewing. Firstly, participants were shown a random sequence of five distinct layouts. These were shown one at a time, and the participants had to request the next layout by clicking a button. The survey provided only a button to move forward, as the viewing time was recorded once for each layout. Figure 5a depicts a sample layout of four tiles shown to a participant.

6.3.2 Phase II: Content ratings. Once the layouts were shown, participants were asked to rate 11 distinct content items, displayed one by one. To this end, two questions were asked: one to rate the content items displayed on the screen; and another to assess the participant's engagement in terms of attention to previously-displayed content items. Participants could rate content items on a scale of 0 to 100, from "least interesting" to "very interesting". Figure 5b shows a sample page, wherein the participant was asked to rate a content item related to weather.

Finally, participants were asked to fill in a demographic questionnaire for statistical purposes. The experiment took, on average, 10 minutes for each participant.

6.4 Data Collection and Preprocessing

The following data were recorded during the experiments: the time spent on each layout shown, measured in seconds; the ratings given to each content item to express interest; and the answers pertaining to the content

Table 2. Model accuracy against the number of layouts shown to the user (N=16).

Layouts	RMSE	R^2
2	0.134	0.560
3	0.147	0.613
4	0.156	0.704
5	0.143	0.787

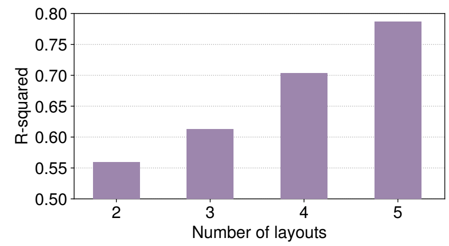


Fig. 6. R^2 as a function of the number of layouts.

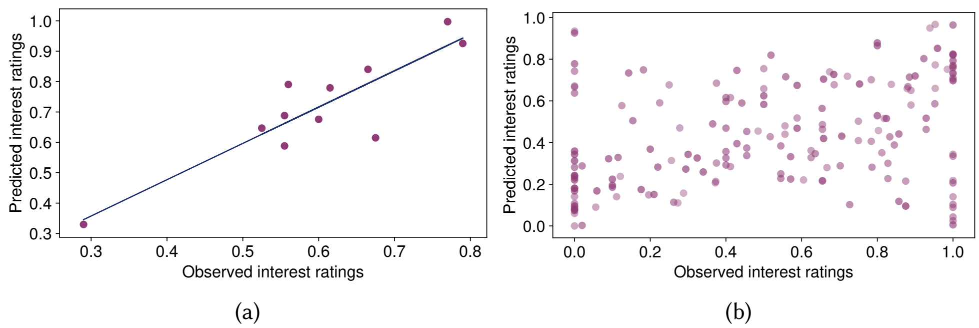


Fig. 7. Observed versus predicted interest found in the empirical user study: (a) for data aggregated by content type, and (b) for individual-level per-content-type data.

items observed earlier. Moreover, the actual content items shown in the experiment were stored to validate participants' answers regarding the content seen. Collected data were checked to verify that all participants in the user study were engaged during the experiments, in terms of long-enough viewing times and consistent answers on content seen in previously-shown layouts. Before fitting the data, the interest ratings were normalized over a smaller scale, and Gaussian noise was added to the viewing-time observations.

6.5 Results

The following reports the results obtained by leveraging the data collected in the user study for inverting the proposed IFT models. Specifically, the considered data included a number of layouts ranging from two to five, leading to estimates of eleven interest values in addition to other parameters (see Section 3). The following focuses on interest values only, as this is the main scope of the work and also because the other parameters are model-specific. The same metrics described in the previous section were considered; interest ratings were normalized to the range 0 to 1 for comparison purposes.

6.5.1 Fitting results. The first set of experiments focuses on fitting accuracy. For simplicity, the obtained results are presented for the tile forager only. Table 2 reports the model accuracy as a function of the number of layouts. The results show that the RMSE slightly increases with the number of layouts, except for the case of five layouts where the RMSE decreases to a value between those obtained for two and three layouts. Instead, the R^2 values always increases with the number of layouts with an almost linear trend, as it could be better seen from Figure 6. These results demonstrate that the proposed model can infer users' interest even when only two layouts are shown; a stronger correlation is achieved when fives are presented, as done in the user study. The following discussion on correlation considers five layouts, as that is the setting providing the best accuracy. Figure 7a shows the correlation between observed versus predicted interest ratings, wherein each data point represents a

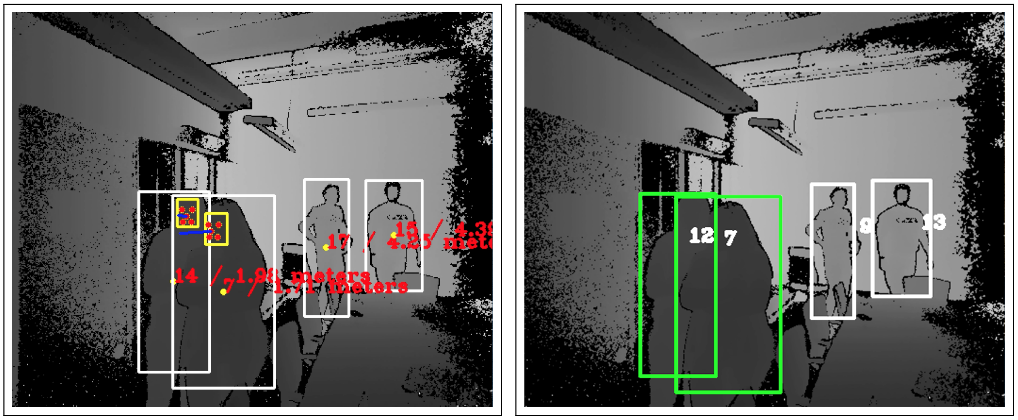
Table 3. Both foraging models achieved very satisfactory accuracy in estimating ground-truth interest data in the empirical user study (N=16). The text in boldface denotes the best values obtained by each model under the considered metrics.

	Tile forager				Display forager			
User	AIC	BIC	RMSE	R^2	AIC	BIC	RMSE	R^2
1	**12.30**	**3.62**	**0.206**	**0.832**	30.16	20.14	0.253	0.684
2	**10.68**	**2.00**	0.191	0.757	16.30	6.28	**0.134**	**0.848**
3	**11.84**	**3.16**	**0.201**	**0.833**	33.41	23.38	0.293	0.733
4	34.83	26.15	0.574	**0.772**	**28.38**	**18.36**	**0.233**	0.753
5	**8.34**	**-0.80**	0.170	0.706	16.49	7.15	**0.143**	**0.795**
6	35.20	26.53	0.583	0.797	**22.08**	**12.06**	**0.175**	**0.873**
7	37.17	28.49	0.638	**0.760**	**31.89**	**21.87**	**0.273**	0.690
8	**24.76**	**16.09**	**0.360**	**0.640**	42.26	32.24	0.451	0.493
9	28.64	19.96	0.433	0.701	**17.12**	**7.10**	**0.140**	**0.755**
10	**10.31**	**1.63**	**0.188**	0.628	46.73	36.71	0.537	**0.824**
11	**10.77**	**2.09**	**0.192**	**0.775**	34.6	24.58	0.309	0.700
12	35.82	27.14	0.600	0.645	**19.37**	**9.35**	**0.155**	**0.839**
13	**5.12**	**-3.55**	**0.148**	**0.897**	32.5	22.48	0.281	0.818
14	**5.58**	**-3.09**	**0.151**	0.883	32.47	22.45	0.281	**0.909**
15	**8.80**	**0.12**	**0.175**	**0.821**	36.48	26.46	0.337	0.638
16	**14.64**	**5.96**	**0.229**	**0.797**	33.22	23.20	0.291	0.782
Mean	**19.37**	**10.66**	0.315	**0.765**	29.59	19.61	**0.268**	0.758
SD	11.93	11.96	0.185	0.083	9.10	9.04	0.112	0.103

content item. The figure clearly shows a strong correlation: the predicted values are close to the observed ones, demonstrating that the model adequately fits. Figure 7b shows the correlation in terms of individual participants, which are represented as data points therein. The results exhibit a similar pattern also in this case.

6.5.2 Comparing the two foraging models. The next set of experiments aims at comparing the display forager and the tile forager. For this purpose, the results are reported for the individual participants, averaged over all the number of layouts in the data. Accordingly, Table 3 presents the accuracy of the two models per user; the best values of the considered metrics are boldfaced for better illustration. First of all, both models achieve high accuracy, with $R^2 > 0.75$ on average. In terms of RMSE, the values obtained by the display forager differ slightly from those for the tile forager, with average values of 0.268 and 0.315, respectively. For a comparison between them, the analysis focuses on ascertaining which model obtains lower values of the AIC and BIC metrics. In this respect, it is clear that the tile forager yields lower values than the display forager for the majority of cases – there are only five cases wherein the display forager performs better than the tile forager. These results are as expected since the display forager searches for more parameters (i.e., the a and b coefficients) than the tile forager (i.e., only α) in addition to the interest ratings. Conversely, the display forager is also more complex; in fact, fitting its parameters takes about twice as much time as for the tile forager. Note that the number of parameters depends both on the number of content items and on the number of layouts.

6.5.3 Insights from collected data. Overall, the time spent viewing the content shown in the layouts ranged between four seconds and two minutes, with an average of 42 seconds. Note that these values are likely to be much higher than those in real scenarios [8, 18, 36], due to the particular conditions considered in the experiment.



(a) (b)

Fig. 8. Sample output of the system prototype for characterizing user behavior. (a) Facial features of the users (represented as red dots) are detected to determine if they are looking at the display or not. (b) This allows to distinguish between users paying attention to the content of the display (in green rectangles) and the rest of the people in the deployment area (in white rectangles).

Nevertheless, they clearly indicate that participants were more engaged with certain layouts than others. The average rating for the content items was 50%, and participants gave similar ratings to all content items in a few cases. However, most of the participants rated at least one of the content items "least interesting" and "very interesting." While responses vary considerably across all the participants, the proposed models can still estimate such values with high accuracy.

7 APPLICATION PROCEDURE FOR INVERSE FORAGING

The results obtained through both simulation and the controlled experiment demonstrated that the models are effective in deriving interest from aggregate viewing times. This is a very important aspect for practical applications. In fact, determining the viewing time of individual tiles is generally cumbersome or impractical, depending on the chosen instrumentation – for instance, gaze trackers generally have a limited range and require prior calibration. In contrast, it is possible to obtain the total time a user spends in perusing a layout (i.e., all the tiles therein) more easily and accurately, as explained next.

Before proceeding further, it is worth noting that inverse foraging as presented here can be employed in scenarios that fulfill the following conditions: (1) at least five distinct multi-content-item layouts are presented over time (2) such that the contents are drawn from a pool of 5–11 content categories, where (3) reliable data are available on how long users spend viewing each layout. The rest of this section focuses on the latter, by introducing a system that reliably measures per-user viewing times with limited instrumentation and no explicit user intervention. The method presented in this article does not distinguish between different users. As a consequence, it could be applied to characterize the audience of pervasive displays on average terms, which is relevant to the considered scenario.

With these considerations in mind, a pervasive display prototype was implemented for inverse foraging. The prototype consists of a client–server architecture wherein the pervasive display client presents layouts produced by a server. It also employs computer vision to track users' viewing durations (Figure 8), without the need for gaze tracking. Specifically, Microsoft Kinect v. 2 was employed to capture depth images of people walking in front of a display, since this system has been widely employed to measure users' attention by detecting their head posture and position [2, 42]. This solution also can recognize facial features of users (i.e., eyes, nose, and mouth) within a two-meter distance range. These features were used to obtain a better estimate of attention. In fact, users too far away from the display cannot distinguish the content shown therein. Moreover, users who are close enough may not be looking at the display.

The prototype was evaluated with a small group of researchers[2] to measure the accuracy of detecting actual views and their duration. Figure 8a presents an example clearly showing two users looking at a display, because their eyes and nose are directed towards the device. The two users behind them, at a distance of four meters from the display, are just considered bystanders. Therefore, the prototype is able not only to determine users and their viewing times but also the number of people present in the deployment area. This is more clearly indicated in Figure 8b, wherein users devoting attention to the display are enclosed within a green rectangle, while the rest of people (ignoring it) with a white rectangle.

8 DISCUSSION

This work has demonstrated the feasibility of using inverse modeling to infer the interest of a user with respect to differing content shown on a display. The approach was evaluated through data collected from a controlled experiment and by means of synthetic data employed to assess parameter recovery. The results demonstrate that, even with fairly few data samples, it is possible to achieve reasonably high accuracy in estimating users' interest.

While the approach has limited scope at present (as discussed in the previous section), it demonstrates several benefits. It is simple and practical to implement – in particular, it does not require user input or depend on eye-tracking devices. Also, the method is applicable in diverse scenarios (e.g., at universities, airports, and shopping malls) since it relies only on viewing-time data, which can be collected in different ways (a survey being only one of them). Practitioners could benefit from this approach to support decisions on selecting certain content options to be shown on a display, since the method provides an indicator of interesting content items, represented as a numerical value.

Some limitations should be noted. Firstly, in the controlled experiment carried out to collect data, the participants were explicitly asked to read the content shown on the screen. Secondly, the task was performed remotely on a laptop or other workstation. Yet, this approach was the most suitable to evaluate the models, since the conditions were similar to those involved in use of a public display. Irrespective of the associated limitations, the data collected can be considered valid for evaluation of the proposed model, because the users explicitly indicated their level of interest in the content shown, using a scale. This technique made it possible to analyze how the predictions differ from the observed data. Hence, the results obtained constitute a promising starting point for further research in this field.

It would be interesting for future work to explore other approaches to improve the modeling's accuracy. One research direction is to add new inputs to the model – for instance, utilizing data collected from tracking devices. To this end, the proposed model could employ information about the number of people standing in front of the screen when certain content items are shown as an indicator of interest. A higher count would indicate that more people are clustered around the display when certain content is presented, thus providing feedback that such content is interesting. Alongside this, the model could make use of the recorded viewing time values as is already done. Furthermore, the proposed prototype could be used to conduct a field study to collect data in a more realistic scenario, thereby establishing ecological validity. Finally, the proposed approach could be complemented by exploring the use of a contextual-bandit technique [32] to optimize the selection of the content items to be shown on the display.

[2]It was not possible to conduct a field study by using this prototype at the time of writing this article due to COVID-related restrictions.

ACKNOWLEDGMENTS

This work was partially supported by: the Academy of Finland under grants number 326346, 332307, 328813 and 318559; the Finnish Center for Artificial Intelligence; and the Finnish Foundation for Technology Promotion. The authors would like to thank Emilio López for his help in implementing the pervasive display prototype and the CSC – IT Center for Science for provisioning the computational resources used in the study.

REFERENCES

[1] Eugene Agichtein, Eric Brill, and Susan Dumais. 2006. Improving Web Search Ranking by Incorporating User Behavior Information. In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (Seattle, WA, USA) *(SIGIR '06)*. New York, NY, USA, 19–26. https://doi.org/10.1145/1148170.1148177

[2] Florian Alt, Andreas Bulling, Lukas Mecke, and Daniel Buschek. 2016. Attention, Please! Comparing Features for Measuring Audience Attention towards Pervasive Displays. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. 823–828.

[3] Stylianos Asteriadis, Kostas Karpouzis, and Stefanos Kollias. 2009. Feature Extraction and Selection for Inferring User Engagement in an HCI Environment. In *Human–Computer Interaction: New Trends*, Julie A. Jacko (Ed.). Springer, Berlin & Heidelberg, Germany, 22–29.

[4] Andreas Bulling. 2016. Pervasive Attentive User Interfaces. *Computer* 49, 1 (2016), 94–98.

[5] Olivier Chapelle and Ya Zhang. 2009. A Dynamic Bayesian Network Click Model for Web Search Ranking. In *Proceedings of the 18th International Conference on World Wide Web* (Madrid, Spain) *(WWW '09)*. New York, NY, USA, 1–10. https://doi.org/10.1145/1526709.1526711

[6] Sarah Clinch, Jason Alexander, and Sven Gehring. 2016. A Survey of Pervasive Displays for Information Presentation. *IEEE Pervasive Computing* 15, 3 (2016), 14–22. https://doi.org/10.1109/MPRV.2016.55

[7] Sunny Consolvo and Miriam Walker. 2003. Using the Experience Sampling Method To Evaluate Ubicomp Applications. *IEEE Pervasive Computing* 2, 2 (2003), 24–31.

[8] Nicholas S. Dalton, Emily Collins, and Paul Marshall. 2015. Display Blindness? Looking Again at the Visibility of Situated Displays Using Eye-Tracking. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. New York, NY, USA, 3889–3898. https://doi.org/10.1145/2702123.2702150

[9] Nigel Davies, Marc Langheinrich, Sarah Clinch, Ivan Elhart, Adrian Friday, Thomas Kubitza, and Bholanathsingh Surajbali. 2014. Personalisation and Privacy in Future Pervasive Display Networks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, ON, Canada) *(CHI '14)*. New York, NY, USA, 2357–2366. https://doi.org/10.1145/2556288.2557287

[10] Antonio Di Ferdinando, Alberto Rosi, Ricardo Lent, Antonio Manzalini, and Franco Zambonelli. 2009. MyAds: A System for Adaptive Pervasive Advertisements. *Pervasive and Mobile Computing* 5, 5 (2009), 385–401.

[11] Yassine Drias and Samir Kechid. 2019. Dynamic Web Information Foraging Using Self-Interested Agents: Application to [a] Scientific Citations Network. *Concurrency and Computation: Practice and Experience* 31, 22 (2019), e4342.

[12] Mary C Dyson. 2004. How physical text layout affects reading from screen. *Behaviour & information technology* 23, 6 (2004), 377–393.

[13] Andreas Fender, David Lindlbauer, Philipp Herholz, Marc Alexa, and Jörg Müller. 2017. HeatSpace: Automatic Placement of Displays by Empirical Analysis of User Behavior. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (Québec City, QC, Canada) *(UIST '17)*. New York, NY, USA, 611–621. https://doi.org/10.1145/3126594.3126621

[14] Gerhard Fischer. 2001. User modeling in human–computer interaction. *User modeling and user-adapted interaction* 11, 1 (2001), 65–86.

[15] Christoph Gebhardt, Antti Oulasvirta, and Otmar Hilliges. 2020. Hierarchical Reinforcement Learning As a Model of Human Task Interleaving. *arXiv preprint arXiv:2001.02122* (2020).

[16] Andrew Gelman, Aki Vehtari, Daniel Simpson, Charles C Margossian, Bob Carpenter, Yuling Yao, Lauren Kennedy, Jonah Gabry, Paul-Christian Bürkner, and Martin Modrák. 2020. Bayesian workflow. *arXiv preprint arXiv:2011.01808* (2020).

[17] Thomas Hills, Mike Jones, and Peter Todd. 2009. Optimal Foraging in Semantic Memory. *Proceedings of the Annual Meeting of the Cognitive Science Society* 31 (2009), 620–625.

[18] Elaine M. Huang, Anna Koster, and Jan Borchers. 2008. Overcoming Assumptions and Uncovering Practices: When Does the Public Really Look at Public Displays?. In *International Conference on Pervasive Computing*. Springer, 228–243.

[19] Giulio Jacucci, Ann Morrison, Gabriela T. Richard, Jari Kleimola, Peter Peltonen, Lorenza Parisi, and Toni Laitinen. 2010. Worlds of Information: Designing for Engagement at a Public Multi-Touch Display. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems*. ACM, 2267–2276. https://doi.org/10.1145/1753326.1753669

[20] Amit Kumar Jaiswal, Haiming Liu, and Ingo Frommholz. 2020. Utilising Information Foraging Theory for User Interaction with Image Query Auto-Completion. In *Advances in Information Retrieval*, Joemon M. Jose, Emine Yilmaz, João Magalhães, Pablo Castells, Nicola Ferro, Mário J. Silva, and Flávio Martins (Eds.). Springer International Publishing, Cham, Switzerland, 666–680.

[21] Jussi P. P. Jokinen, Sayan Sarcar, Antti Oulasvirta, Chaklam Silpasuwanchai, Zhenxin Wang, and Xiangshi Ren. 2017. Modelling Learning of New Keyboard Layouts. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, CO, USA) *(CHI*

'17). New York, NY, USA, 4203–4215. https://doi.org/10.1145/3025453.3025580

[22] Rui José, Jorge Cardoso, Florian Alt, Sarah Clinch, and Nigel Davies. 2013. Mobile Applications for Open Display Networks: Common Design Considerations. In *Proceedings of the 2nd ACM International Symposium on Pervasive Displays* (Mountain View, California) *(PerDis '13)*. New York, NY, USA, 97–102. https://doi.org/10.1145/2491568.2491590

[23] Antti Kangasrääsiö, Kumaripaba Athukorala, Andrew Howes, Jukka Corander, Samuel Kaski, and Antti Oulasvirta. 2017. Inferring Cognitive Models from Data Using Approximate Bayesian Computation. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, CO, USA) *(CHI '17)*. New York, NY, USA, 1295–1306. https://doi.org/10.1145/3025453.3025576

[24] Antti Kangasrääsiö, Jussi P. P. Jokinen, Antti Oulasvirta, Andrew Howes, and Samuel Kaski. 2019. Parameter Inference for Computational Cognitive Models with Approximate Bayesian Computation. *Cognitive Science* 43, 6 (2019), e12738.

[25] Mary Beth Kery, Bonnie E. John, Patrick O'Flaherty, Amber Horvath, and Brad A. Myers. 2019. Towards Effective Foraging by Data Scientists To Find Past Analysis Choices. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland) *(CHI '19)*. New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300322

[26] Mohamed Khamis, Axel Hoesl, Alexander Klimczak, Martin Reiss, Florian Alt, and Andreas Bulling. 2017. EyeScout: Active Eye Tracking for Position and Movement Independent Gaze Interaction with Large Public Displays. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (Quebec City, QC, Canada) *(UIST '17)*. New York, NY, USA, 155–166. https://doi.org/10.1145/3126594.3126630

[27] Thomas Kubitza, Sarah Clinch, Nigel Davies, and Marc Langheinrich. 2013. Using Mobile Devices To Personalize Pervasive Displays. *ACM SIGMOBILE Mobile Computing and Communications Review* 16, 4 (2013), 26–27.

[28] Sandeep Kaur Kuttal, Anita Sarma, Margaret Burnett, Gregg Rothermel, Ian Koeppe, and Brooke Shepherd. 2019. How End-User Programmers Debug Visual Web-Based Programs: An Information Foraging Theory Perspective. *Journal of Computer Languages* 53 (2019), 22–37. https://doi.org/10.1016/j.cola.2019.04.003

[29] Jeffrey C Lagarias, James A Reeds, Margaret H Wright, and Paul E Wright. 1998. Convergence properties of the Nelder–Mead simplex method in low dimensions. *SIAM Journal on optimization* 9, 1 (1998), 112–147.

[30] Joseph Lawrance, Rachel Bellamy, Margaret Burnett, and Kyle Rector. 2008. Using Information Scent To Model the Dynamic Foraging Behavior of Programmers in Maintenance Tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. New York, NY, USA, 1323–1332. https://doi.org/10.1145/1357054.1357261

[31] J. Lawrance, C. Bogart, M. Burnett, R. Bellamy, K. Rector, and S. D. Fleming. 2013. How Programmers Debug, Revisited: An Information Foraging Theory Perspective. *IEEE Transactions on Software Engineering* 39, 2 (2013), 197–215.

[32] Lihong Li, Wei Chu, John Langford, and Robert E. Schapire. 2010. A Contextual-Bandit Approach to Personalized News Article Recommendation. In *Proceedings of the 19th International Conference on World Wide Web* (Raleigh, NC, USA) *(WWW '10)*. New York, NY, USA, 661–670. https://doi.org/10.1145/1772690.1772758

[33] Haiming Liu, Paul Mulholland, Dawei Song, Victoria Uren, and Stefan Rüger. 2010. Applying Information Foraging Theory To Understand User Interaction with Content-Based Image Retrieval. In *Proceedings of the Third Symposium on Information Interaction in Context* (New Brunswick, NJ, USA) *(IIiX '10)*. New York, NY, USA, 135–144. https://doi.org/10.1145/1840784.1840805

[34] Jiahui Liu, Peter Dolan, and Elin Rønby Pedersen. 2010. Personalized News Recommendation Based on Click Behavior. In *Proceedings of the 15th International Conference on Intelligent User Interfaces* (Hong Kong, China) *(IUI '10)*. New York, NY, USA, 31–40. https://doi.org/10.1145/1719970.1719976

[35] Hongyu Lu, Min Zhang, and Shaoping Ma. 2018. Between Clicks and Satisfaction: Study on Multi-Phase User Preferences and Satisfaction for Online News Reading. In *The 41st International ACM SIGIR Conference on Research and Development in Information Retrieval* (Ann Arbor, MI, USA) *(SIGIR '18)*. New York, NY, USA, 435–444. https://doi.org/10.1145/3209978.3210007

[36] Ville Mäkelä, Rivu Radiah, Saleh Alsherif, Mohamed Khamis, Chong Xiao, Lisa Borchert, Albrecht Schmidt, and Florian Alt. 2020. Virtual Field Studies: Conducting Studies on Public Displays in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. New York, NY, USA, 1–15. https://doi.org/10.1145/3313831.3376796

[37] Nemanja Memarovic, Sarah Clinch, and Florian Alt. 2015. Understanding Display Blindness in Future Display Deployments. In *Proceedings of the 4th International Symposium on Pervasive Displays*. 7–14.

[38] Mateusz Mikusz, Anastasios Noulas, Nigel Davies, Sarah Clinch, and Adrian Friday. 2016. Next Generation Physical Analytics for Digital Signage. In *Proceedings of the 3rd International on Workshop on Physical Analytics* (Singapore, Singapore) *(WPA '16)*. New York, NY, USA, 19–24. https://doi.org/10.1145/2935651.2935658

[39] Maria L. Montoya Freire, Dominic Potts, Niraj Dayama, Antti Oulasvirta, and Mario Di Francesco. 2019. Foraging-based Optimization of Pervasive Displays. *Pervasive and Mobile Computing* 55 (April 2019), 45–58. https://doi.org/10.1016/j.pmcj.2019.02.008

[40] Jörg Müller and Antonio Krüger. 2007. User Profiling for Generating Bids in Digital Signage Advertising Auctions. *UBIDEUM 2007* 46 (2007).

[41] Jörg Müller, Dennis Wilmsmann, Juliane Exeler, Markus Buzeck, Albrecht Schmidt, Tim Jay, and Antonio Krüger. 2009. Display Blindness: The Effect of Expectations on Attention towards Digital Signage. In *International Conference on Pervasive Computing*. Springer, 1–8.

[42] Wolfgang Narzt. 2017. A Comparison of Attention Estimation Techniques in a Public Display Scenario. In *HCI in Business, Government and Organizations: Interacting with Information Systems*, Fiona Fui-Hoon Nah and Chuan-Hoo Tan (Eds.). Springer International Publishing, Cham, Switzerland, 338–353.

[43] Allen Newell. 1982. The knowledge level. *Artificial intelligence* 18, 1 (1982), 87–127.

[44] Antti Oulasvirta. 2019. It's time to rediscover HCI models. *interactions* 26, 4 (2019), 52–56.

[45] Jake Patterson and Sarah Clinch. 2018. SlideTalk: Encouraging User Engagement with Slideshow Displays. In *Proceedings of the 7th ACM International Symposium on Pervasive Displays* (Munich, Germany) *(PerDis '18)*. New York, NY, USA, Article 4, 7 pages. https://doi.org/10.1145/3205873.3205883

[46] Peter Peltonen, Esko Kurvinen, Antti Salovaara, Giulio Jacucci, Tommi Ilmonen, John Evans, Antti Oulasvirta, and Petri Saarikko. 2008. It's Mine, Don't Touch! Interactions at a Large Multi-Touch Display in a City Centre. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. New York, NY, USA, 1285–1294. https://doi.org/10.1145/1357054.1357255

[47] Peter Pirolli. 2007. *Information Foraging Theory: Adaptive Interaction with Information.* Oxford University Press.

[48] Peter Pirolli and Wai-Tat Fu. 2003. SNIF-ACT: A Model of Information Foraging on the World Wide Web. *Lecture Notes in Artificial Intelligence* 2702 (2003), 45–54.

[49] Michael JD Powell. 1964. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *The computer journal* 7, 2 (1964), 155–162.

[50] Kenneth V. Price. 2013. *Differential Evolution.* Springer, Berlin & Heidelberg, Germany, 187–214. https://doi.org/10.1007/978-3-642-30504-7_8

[51] Feng Qiu and Junghoo Cho. 2006. Automatic Identification of User Interest for Personalized Search. In *Proceedings of the 15th International Conference on World Wide Web* (Edinburgh, Scotland) *(WWW '06)*. New York, NY, USA, 727–736. https://doi.org/10.1145/1135777.1135883

[52] Gianluca Schiavo, Eleonora Mencarini, Kevin B. A. Vovard, and Massimo Zancanaro. 2013. Sensing and Reacting to Users' Interest: An Adaptive Public Display. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (Paris, France) *(CHI EA '13)*. New York, NY, USA, 1545–1550. https://doi.org/10.1145/2468356.2468632

[53] Tobias Schnabel, Paul N. Bennett, and Thorsten Joachims. 2019. Shaping Feedback Data in Recommender Systems with Interventions Based on Information Foraging Theory. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining* (Melbourne VIC, Australia) *(WSDM '19)*. New York, NY, USA, 546–554. https://doi.org/10.1145/3289600.3290974

[54] Mohit Sethi, Elena Oat, Mario Di Francesco, and Tuomas Aura. 2014. Secure Bootstrapping of Cloud-Managed Ubiquitous Displays. In *The 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2014)*. 739–750. https://doi.org/10.1145/2632048.2632049

[55] Matthew Sharifi, Terry Payne, and Esther David. 2006. Public Display Advertising Based on Bluetooth Device Presence. In *Proceedings of the Workshop Mobile Interaction with the Real World (MIRW 2006)*. 52–55.

[56] Si Shen, Botao Hu, Weizhu Chen, and Qiang Yang. 2012. Personalized Click Model through Collaborative Filtering. In *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining* (Seattle, WA, USA) *(WSDM '12)*. New York, NY, USA, 323–332. https://doi.org/10.1145/2124295.2124336

[57] Yusuke Sugano, Xucong Zhang, and Andreas Bulling. 2016. AggreGaze: Collective Estimation of Audience Attention on Public Displays. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. New York, NY, USA, 821–831. https://doi.org/10.1145/2984511.2984536

[58] Albert Tarantola. 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation.* SIAM.

[59] Peter M. Todd and Thomas T. Hills. 2020. Foraging in Mind. *Current Directions in Psychological Science* (2020), 0963721420915861.

[60] Pooja Upadhyay. 2020. Comparing Non-Visual and Visual Information Foraging on the Web. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. New York, NY, USA, 1–8. https://doi.org/10.1145/3334480.3383025

[61] Mélodie Vidal, Andreas Bulling, and Hans Gellersen. 2013. Pursuits: Spontaneous Interaction with Displays Based on Smooth Pursuit Eye Movement and Moving Targets. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Zurich, Switzerland) *(UbiComp '13)*. New York, NY, USA, 439–448. https://doi.org/10.1145/2493432.2493477

[62] Robert C. Wilson and Anne G. E. Collins. 2019. Ten Simple Rules for the Computational Modeling of Behavioral Data. *eLife* 8 (Nov. 2019), e49547. https://doi.org/10.7554/eLife.49547

[63] Xing Yi, Liangjie Hong, Erheng Zhong, Nanthan Nan Liu, and Suju Rajan. 2014. Beyond Clicks: Dwell Time for Personalization. In *Proceedings of the 8th ACM Conference on Recommender Systems* (Foster City, CA, USA) *(RecSys '14)*. New York, NY, USA, 113–120. https://doi.org/10.1145/2645710.2645724

[64] Yanxia Zhang, Ming Ki Chong, Jörg Müller, Andreas Bulling, and Hans Gellersen. 2015. Eye Tracking for Public Displays in the Wild. *Personal and Ubiquitous Computing* 19, 5–6 (2015), 967–981.

[65] Guorui Zhou, Na Mou, Ying Fan, Qi Pi, Weijie Bian, Chang Zhou, Xiaoqiang Zhu, and Kun Gai. 2019. Deep Interest Evolution Network for Click-Through Rate Prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 5941–5948.