Semantic Pervasive Advertising*

Lorenzo Carrara[1], Giorgio Orsi[2], and Letizia Tanca[3]

[1] Cubica S.r.l., Italy
lorenzo.carrara@cubica.eu
[2] Department of Computer Science, University of Oxford, UK
giorgio.orsi@cs.ox.ac.uk
[3] Dipartimento di Elettronica e Informazione, Politecnico di Milano, Italy
tanca@elet.polimi.it

Abstract. Pervasive advertising targets consumers on-the-move with ads displayed on their mobile devices. As for web advertising, ads are distributed by embedding them into websites and apps, easily flooding consumers with a large number of uninteresting offers. As the pervasive setting amplifies traditional issues such as targeting, cost, and privacy, we argue the need for a new perspective on the problem. We introduce PervADs, a privacy-preserving, user-centric, and pervasive ads-distribution platform which uses semantic technologies to reason about the consumer's context and the intended target of the ads.

1 Introduction

Advertising is a form of business marketing communication which produces awareness of an audience w.r.t. an offer about a product or service. The communication occurs between *advertisers* and the *consumers* over one or more *media* (e.g., TV) by means of *ads* whose goal is to imprint the product in the mind of the consumer.

Targeting. The distribution of ads from multiple advertisers via mass communication channels such as TV, newspapers, and the Web, easily floods consumers with uninteresting commercial offers. As a result, the ads tend to be perceived as background noise, harming the overall effectiveness of the advertising campaign. As a consequence: (i) the ads will be ineffective outside the target audience and, (ii) even when they reach the target, the consumer might not be interested in the offers in that specific moment.

Cost. Advertising campaigns have a considerable economical impact on businesses. A single 30-seconds commercial on US TV channels cost in average 100K$, it doubles if it is shown during popular TV shows, and reaches peaks of 2M$ during special events such as the SuperBowl. The cost is prohibitive for the majority of small and medium businesses (SMBs) that then turned their attention to cheaper channels, e.g., radio instead of TV commercials, thus reducing the effectiveness of the campaign, or by

* G. Orsi acknowledges the Oxford Martin Schools grant no. LC0910-019 and the ERC grant DIADEM no. 246858. L. Tanca acknowledges the Lombardy region project Green Move. Full technical report available at: http://www.cs.ox.ac.uk/files/4735/RR-11-07.pdf. PervADs's code is available open-source at: http://code.google.com/p/pervads/.

W. Faber and D. Lembo (Eds.): RR 2013, LNCS 7994, pp. 216–222, 2013.

turning to the web that democratized this field by lowering the cost of advertising and by increasing its effectiveness [14].

Context Awareness. The Web enabled more sophisticated and interactive forms of advertising that allow, e.g., profile-based crafting of the ads. In the same direction, pervasive technologies enabled geo-localised mobile advertising that targets consumers during their everyday activity and that is typically realised by placing contextual ads into web pages and mobile applications[1]. Due to its potential, *pervasive advertising* is expected to generate billions of dollars of revenues in the near future [5]; however, a study by Sala et. Al [17] has shown that not all contextual information influences its effectiveness, e.g., activity information has very limited use for targeting purposes.

Privacy. Private user data are necessary to properly target the ads and are collected by advertisers in exchange for certain services (e.g., social networking). These policies have a great impact on privacy since users may disclose (often unconsciously) personal data which can be later used in uncontrolled ways. In the pervasive setting, privacy-related issues can become even more serious since it is possible to track user locations through location-aware services (e.g., Facebook Places). Aggressive behaviour in advertising strategies lead the European Commission to publish regulations such as the EU Directive 2002/58/EC also known as E-privacy Directive. Since its introduction, European online advertising suffered a 65% loss in revenues [7], a situation that created problems not only to advertising companies but also to businesses.

Challenges. The pervasive setting amplifies traditional challenges of advertising such as targeting and privacy and can make it more expensive than online advertising, e.g., the reduced size of the screens makes the ads-auctions more competitive. We argue that, in order to be effective, pervasive advertising has to deal with targeting, privacy and cost in an integrated fashion, moving away from the traditional advertiser-centric approaches adopted so far. Semantic technologies are key to enabling this shift, since pervasive advertising often requires to reason about complex objects such as product descriptions, user profiles and contextual information.

Contributions. We propose PervADs: an ads/coupon distribution platform for pervasive environments providing: *(i)* a rich semantic formalism for the description of pervasive ads (*PAD*s), *(ii)* a novel architecture for their pervasive distribution, and *(iii)* a client-side, ads-filtering mechanism based on context matching, ensuring the privacy of personal data. PervADs enables one-to-one interaction between businesses and consumers without intermediate third-party entities. The PADs published by businesses are locally and privately filtered on the devices of potential customers making very hard for advertisers to collect personal consumer's data. PervADs also provides businesses with an autonomous and inexpensive advertising infrastructure enabling a fine-grained monitoring of the performance of the advertising campaign.

2 Related Work

Computational advertising is concerned with problems in automated advertising. An example is *sponsored search* where search results related to paying businesses are

[1] See e.g., Google AdMob.

shown in privileged ranking positions. Our interest lies in *contextual advertising* and *behavioural targeting*. The former places ads on the basis of the situational interests and activities of the consumer; the latter exploits observations about their behaviour.

Context awareness and *semantics* are key factors in computational advertising as acknowledged by most of the current approaches as [15,10,4,16]. The context is commonly represented as a combination of user location and profile represented as key-value pairs or as an organised taxonomy of keywords. As a consequence, users and ads can be matched using techniques coming from information retrieval, such as the vector-space model and the cosine similarity. However, various studies in computational advertising demonstrate how keyword matching often yields poor performance in targeting the ads since plain keywords carry poor semantics. Context can also be inferred from the *activity* of the users (see, e.g., [8]). However, it has been proven that activity-related information has little impact on the effectiveness of targeting [17]. More reliable approaches are those making use of background knowledge (see, e.g., [11,2]) possibly coupled with probabilistic models [18] to improve the targeting of the ads.

There is a general agreement about the need of enforcing strict privacy policies in context-aware pervasive computing environments [14,3]. According to the principles outlined in [14], PervADs enforces user privacy by hiding personal information from advertisers while allowing the measurement of the performance metrics about the ads. The client-side matching of ads is what differentiates PervADs from other e-couponing systems like AdNext [9], MyAds [6], AroundMe, and ShopKick[2]; in fact, it removes the need for mediators and enables wide access to personal data to improve targeting.

3 The PervADs Approach

PervADs enables a direct communication between the *advertiser* and the *consumer*. The former is equipped with a customized WiFi router which is used to craft and emit PADs. The latter carries a mobile device which can receive, filter and display PADs on the basis of the consumer's context. Once the PADs have been received, only the coupon is displayed and the user can inspect the details of the ads. The infrastructure is based on wireless technologies without need for accessing the Web and thus enabling free-of-charge communication between the user and the advertiser. Typical tasks include:

1. *Profile and search queries.* Users input personal data used by the receiver to compute the relevance of the PADs. These data are organized in two distinct sets: (i) *Consumer profile*, not bound to a specific user task, i.e., describing time- and situation-independent facts such as dietary preferences. (ii) *Search queries*, describing contingent needs and asynchronously activated by the user. As an example, the user might configure a query to scan for lunch-offers in restaurants located within 500 metres from the user current position at lunch time.

2. *Gathering.* PADs can be collected either by (i) *synchronous scanning*, i.e., the user actively asks for an update of the PADs, possibly specifying one or more search queries, or by (ii) *asynchronous scanning*, i.e., the user activates a periodic and passive refresh of the PADs. In both cases, the application filters them on the basis of the contextual data and notifies the user only when new relevant ads are received.

[2] http://www.shopkick.com/.

3. *Interaction.* The user interacts with the human-readable content of the PADs. Advertisers interact with the administration application installed on their PervADs hotspots and supporting the following administration tasks:

1. *PAD creation and management.* Advertisers define the advertisement context, and then describe the ads details. Each PAD can describe multiple offers but is uniquely associated with a target context.

2. *Performance analysis*: A PAD can be monitored to determine its effectiveness. Performance analysis can be accomplished by standard means used in web advertising. For example, PADs can notify the hotspot whenever an ads is accepted or rejected.

Context Modelling. The target of the ads is defined by specifying in which context the consumer should be in order to receive the offer. In PervADs, this is modelled using the Context-Dimension Ontology (CDO) [12] that has proven to be an effective tool for context-aware data access [13] and that encompasses the work of [1].

The core elements in a CDO are the context dimensions, e.g., `interest-topic`, that are then assigned to values, e.g., `food`. The CDO allows hierarchical dimension assignments e.g., an assignment `region=chinese` is a special case of an assignment `interest-topic=food`. Another useful feature is the possibility to decorate the values with parameters describing specific facets. As an example, the assignment `location=nearby` can be parametrized with a variable `$radius` specifying the radius we are referring to, e.g., `within 50 metres`.

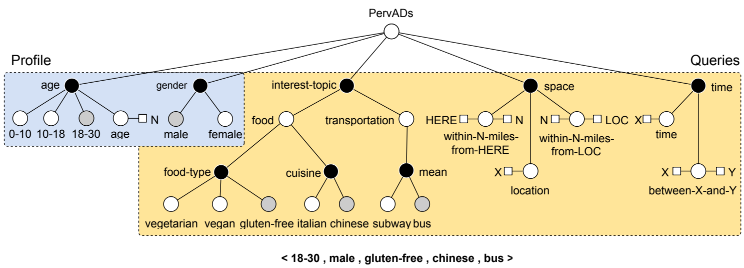
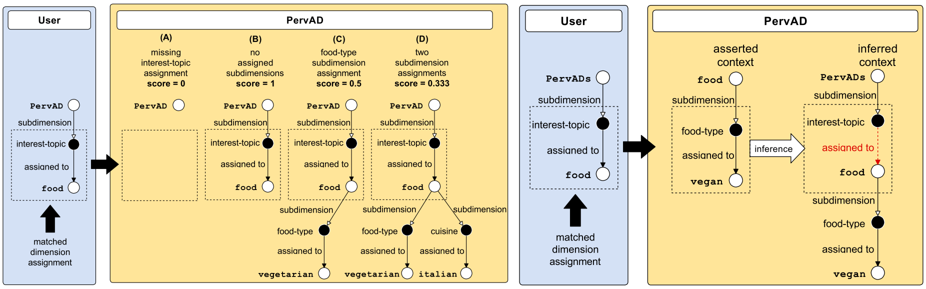


Fig. 1. The PervADs context schema (excerpt)

The context is a combination of a *context schema* and a *context instance*. A context-schema is a tree (see Figure 1), where dimensions (resp. values) are represented as black (resp. white) nodes while parameters are represented as square nodes. A dimension assignment is represented as an edge between a white node and a black node while a sub-dimension is represented as an edge between a white and a black node. A context instance is a set of dimension assignments organized in a tree hierarchy and that is compliant with the context schema.

A possible context instance can be graphically represented as a selection of white nodes (with associated parameters) such that, for each dimension, only one white node belongs to the set. Consider again Figure 1; the darkened nodes identify the context of 18-30 years-old man interested in gluten-free chinese restaurants reachable by bus. In practical cases, each context usually contains between 1 and 10 dimension assignments.

PAD Filtering. Selection and ranking of PADs is based on a matching algorithm that compares the context of the advertisement and that of the user. The matching algorithm takes as input two context instances and produces as output a similarity measure called *context similarity*. The algorithm proceeds by first computing the similarity between each assignment in the user's context and the assignments in the PAD's context.



(a) (b)

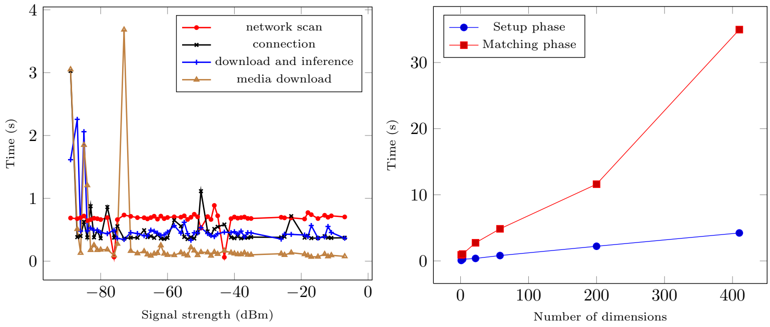
Fig. 2. Context matching and inference

Figure 2a shows some matching examples. In situation (A) the context of the PAD does not define an assignment for the `interest-topic` dimension, so the dimension similarity for `interest-topic` is zero and, in turn, the context similarity is also zero. In the second case (B) the PAD's context defines an assignment to the same value of the user's context and does not assign any value to its sub-dimensions, thus, its dimension-similarity is 1. Cases (C) and (D) respectively add to case (B) one and two assignments to sub-dimensions. The dimension similarity in these cases is 1 for (c) and 2 for (D) resulting in a context similarity of 0.5 and 0.33 respectively.

Since the contexts might have been wrongly or partially specified, PervADs uses reasoning in order to: (i) *validate* a context instance and (ii) *complete* an under-specified context instance. Figure 2b shows an example of context inference where the context instance of the user has the value `food` for the dimension `interest-topic`, while the context of the PAD carries the value `vegan` for the dimension `food-type`. In this case the assignment `(food,interest-topic)` is automatically inferred due to the constraints of the context schema of Figure 1.

System Implementation. PervADs hotspots come as an extension to OpenWRT[3], running on a Linksys WRT54GL WiFi router. The PervADs server application is a LuCI[4] extension written in Lua and based on the uHTTPd package. The PervADs client has been developed as an Android application based on AndroJena[5] ontology modelling and querying. PervADs have been tested using a Samsung Galaxy S i9000 smartphone with 1GHz processor and 64Mb of heap size. The dataset consists of six different context schemata based on taxonomies of categories from UNSPSC and the Yellow Pages.

[3] `http://www.openwrt.org`.
[4] `http://luci.subsignal.org/`.
[5] `http://code.google.com/p/androjena/`.



(a) (b)

Fig. 3. Discovery and matching performance

Figure 3 shows the time required to discover, download and analyse the PADw.r.t. the signal's strength.

References

1. Bolchini, C., Quintarelli, E., Tanca, L.: Carve: Context-aware automatic view definition over relational databases. Inf. Syst. 38(1), 45–67 (2013)
2. Broder, A., Fontoura, M., Josifovski, V., Riedel, L.: A semantic approach to contextual advertising. In: SIGIR, pp. 559–566 (2007)
3. Bulander, R., Decker, M., Schiefer, G., Kolmel, B.: Comparison of different approaches for mobile advertising. In: WMCS, pp. 174–182 (2005)
4. de Castro, J.E., Shimakawa, H.: Mobile advertisement system utilizing user's contextual information. In: MDM, pp. 91–95 (2006)
5. Wireless Federation. Wireless news, analysis: are mobile ads the next gold rush? (2007)
6. Di Ferdinando, A., Rosi, A., Lent, R., Manzalini, A., Zambonelli, F.: MyAds: A system for adaptive pervasive advertisements. Pervasive Mob. Comput. 5(5), 385–401 (2009)
7. Goldfarb, A., Tucker, C.E.: Privacy regulation and online advertising. Management Science 57(1), 57–71 (2011)
8. Haddadi, H., Hui, P., Brown, I.: MobiAd: private and scalable mobile advertising. In: MobiArch, pp. 33–38 (2010)
9. Kim, B., Ha, J., Lee, S., Kang, S., Lee, Y., Rhee, Y., Nachman, L., Song, J.: AdNext: a visit-pattern-aware mobile advertising system for urban commercial complexes. In: HotMobile, pp. 7–12 (2011)
10. Kurkovsky, S., Harihar, K.: Using ubiquitous computing in interactive mobile marketing. Personal Ubiquitous Comput. 10(4), 227–240 (2006)
11. Mirizzi, R., Ragone, A., di Noia, T., di Sciascio, E.: Semantic tags generation and retrieval for online advertising. In: CIKM, pp. 1089–1098 (2010)
12. Orsi, G.: Context Based Querying of Dynamic and Heterogeneous Information Sources. PhD thesis, Politecnico di Milano (2011)
13. Orsi, G., Zimeo, E., Tanca, L.: Context aware, keyword based selection of natural language query patterns. In: EDBT, pp. 189–200 (2011)
14. Ranganathan, A., Campbell, R.H.: Advertising in a pervasive computing environment. In: WMC, pp. 10–14 (2002)

15. Rashid, O., Coulton, P., Edwards, R.: Providing location based information/advertising for existing mobile phone users. Personal Ubiquitous Comput. 12(1), 3–10 (2008)
16. Ratsimor, O., Finin, T., Joshi, A., Yesha, Y.: eNcentive: a framework for intelligent marketing in mobile peer-to-peer environments. In: ICEC, pp. 87–94 (2003)
17. Sala, M.C., Partridge, K., Jacobson, L., "Bo" Begole, J.: An exploration into activity-informed physical advertising using PEST. In: LaMarca, A., Langheinrich, M., Truong, K.N. (eds.) Pervasive 2007. LNCS, vol. 4480, pp. 73–90. Springer, Heidelberg (2007)
18. Wu, X., Yan, J., Liu, N., Yan, S., Chen, Y., Chen, Z.: Probabilistic latent semantic user segmentation for behavioral targeted advertising. In: ADKDD, pp. 10–17 (2009)