Targeted Advertising on the Handset: Privacy and Security Challenges

Hamed Haddadi, Pan Hui, Tristan Henderson and Ian Brown

Abstract Online advertising is currently a rich source of revenue for many Internet giants. With the ever-increasing number of smart phones, there is a fertile market for personalised and localised advertising. A key benefit of using mobile phones is to take advantage of the significant amount of information on phones — such as locations of interest to the user — in order to provide personalised advertisements. Preservation of user privacy, however, is essential for successful deployment of such a system. In this chapter we provide an overview of existing advertising systems and privacy concerns on mobile phones, in addition to a system, MobiAd, which includes protocols for scalable local advertisement download and privacy-aware click report dissemination. In the final section of this chapter we describe some of the security mechanisms used in detecting click-through fraud, and techniques that can be used to ensure that the extra privacy protections of MobiAd are not abused to defraud advertisers.

Key words: Internet advertising, Profiling, Privacy, Mobility

Hamed Haddadi
University of London, UK, e-mail: `hamed@ee.ucl.ac.uk`

Pan Hui
Deutsche Telekom Laboratories, Germany, e-mail: `Pan.Hui@telekom.de`

Tristan Henderson
University of St Andrews, UK, e-mail: `tristan@cs.st-andrews.ac.uk`

Ian Brown
Oxford Internet Institute, UK, e-mail: `ian.brown@oii.ox.ac.uk`

1 Introduction

Advertising is one of the largest revenue sources of many Internet giants. Targeted and personalised advertisements, provided by advertising brokers such as Google and Microsoft, are displayed on designated advertisement slots on websites that in return receive payment from the advertising network. Google's advertising revenue in 2010 was over $28 billion and is only expected to increase over time.[1]

The mobile phone advertising market is also becoming increasingly significant. There are currently over 3 billion mobile phone subscribers in the world. Surveys from Gartner and Telsyte group suggest that nearly a third of these are using smart phones, with the smart phone market increasing at a rate of nearly 50% last year (Cozza et al, 2008). With modern smart phones having 3G and wireless connectivity, GPS and Wi-Fi localisation capabilities, a wide range of social networking applications and web-browsing abilities on large touch LCD displays, there is a fertile market for targeted and personalised advertising. Naturally, handset manufacturers have recently launched a series of advertising platforms which leverage the users' choice of websites, activities, music and social activities to present them with targeted advertisements.

There are a large number of technical, legal and user-related obstacles to overcome on the path to a successful mobile advertising strategy. The use of sensitive, personal information kept on the phones can raise privacy concerns, and successful and accurate profiling and personalisation of advertisements will depend strongly on advertising networks assuaging consumers' privacy concerns over targeted advertisements (Turow and Hennessy, 2007). Mobile phones in general have also less bandwidth, processing power and screen size when compared to ordinary computers. Hence any advertisements must be smaller, downloaded less frequently and have low processing requirements. The profiling tasks must also require limited computation and storage access to preserve battery life.

This chapter describes MobiAd, a system for personalised, localised and targeted advertising on smart phones. Utilising the rich set of information available on the phone, MobiAd presents the user with local advertisements in a privacy-preserving manner. Advertisements are selected by the phone from the pool of advertisements which are broadcast on the local mobile base station or received from local Wi-Fi hotspots. In this manner, the user only needs to download advertisements which are relevant to his interests, and are for items and services in his locality. Information about advertisement views and clicks are then encrypted and sent to the advertisement channel via other mobile phones and intermittent Wi-Fi hotspots, in a delay-tolerant manner. In this system, other nodes and the network operator cannot discover which advertisements were viewed. Likewise, the advertisement provider cannot determine which users viewed which advertisements and only receives aggregate information. MobiAd allows businesses, both local and global, to target users narrowly and directly, without compromising users' privacy. It also improves

---

[1] http://investor.google.com/fin_data.html

the scalability of advertisement distribution by using a local broadcast frequency with geo-targeted advertisements.

In this chapter we also discuss the security issues surrounding online advertising, such as click fraud, where a weblog publisher continuously clicks on the advertisements displayed on his own website in order to make revenue. Detecting click fraud is a relatively new area of research. In its simplest form, the broker can perform threshold-based detection. If a web page is receiving a high number of clicks from the same IP address in a short interval, these clicks can be flagged as fraud. The detection gets complicated if the clickers are behind proxies or globally distributed. We present *Bluff Ads* (Haddadi, 2010), a set of advertisements which are designed to be detected and clicked only by machines, or poorly trained click-fraud work force. These advertisements are targeted at the same audience profile as the other advertisement groups, however their displayed text is totally unrelated to the user profile. Hence they should not be clicked on by the benign user. This simple set of advertisements, mixed with ordinary advertisements, work as a litmus test, or a "CAPTCHA" for the user's legitimacy. If a high number of Bluff ads are clicked, the user is deemed to be flagged as suspicious. Another form of Bluff ads is a set which contain specialised text but they are not targeted to a specific profile and are randomly displayed. This group helps in detecting click-fraud when the botnet builds up a fake profile to harvest relevant ads.

2 Advertising and Privacy

Despite the fertile market for advertising, there are not many dedicated advertising networks for mobile phones. There are a few services for serving advertisements on mobile websites. For instance, AdMob is a service which provides advertisements for more than 15,000 mobile Web sites and applications around the world. AdMob stores and analyses the data from every advertisement request, impression, and click and uses this to optimise advertisement-matching in its network (AdMob, 2010). However the methods used are in no way privacy-aware or localised. This limits the scalability of the system as advertisements have to be served individually at the time of browsing. This is not an issue in general for desktops, but on a mobile phone numerous HTTP connections could slow down the browsing experience.

Adnostic (Toubiana et al, 2010) and Privad (Guha et al, 2009) are also newly-proposed private advertising systems for ordinary browsers. They work on the basis of downloading all of the relevant advertisements offline and showing them at appropriate times. The core ideas of these systems are similar to MobiAd from a privacy perspective. Operation in a mobile environment, however, brings a range of challenges on dissemination of advertisements, capturing reports and scalability. We have attempted to address these issues by using a range of solutions such as DTN for report collection and 3G broadcast channel for advertisement dissemination. MobiAd is also resistant to collusion between advertisers and network opera-

tors, as the DTN anonymisation strategy would prevent the origin of the clicks being easily traced.

Recently, Apple and Microsoft have also entered the mobile advertising market. Apple launched the *iAd* service[2], on which they will perform a range of standard targeting options including demographics, application preferences, music and movies choice and location. All of this information will be kept by Apple and will be used to target advertisements to relevant customers. Google use advertisements throughout their Android smartphone operating system, while Microsoft also envisage a similar service on the Windows Mobile platform. Having such detailed profile information at a content provider or handset provider's disposal is a clear threat to users' privacy. There are some suggested solutions for managing cookies and trackers, but they usually require detailed analysis of the cost-benefit trade-offs Freudiger et al (2009).

The concept of mobile or pervasive advertising has been researched for several years, e.g., (Ranganathan and Campbell, 2002). Mobile advertising systems have been built and studied using existing technologies such as Bluetooth (Aalto et al, 2004) to test their viability. User studies have also been carried out on simple mobile advertising scenarios, such as the sale of ringtones (Merisavo et al, 2006), that indicate that users who are exposed to such advertising do indeed purchase the advertised services. M-system has also been introduced by Komulainen et al (2006) as a permission based advertising for the use of local retailers and consumers. In this system, service provider hosted the system and gathered and updated databases of consumers. The m- advertisers or their advertising agencies created and sent the m-ads by using the advertising tool. However the profile information was provided by consumer and kept up-to-date in the central database. They found that almost all users find privacy a concern in this system which could potentially be compensated by monetary or entertainment value. However, retailers naturally had great interest in use of the system.

But despite this interest in the area, we still lack high-quality data about how, where, and when consumers are willing to allow mobile advertising, or indeed if they would be willing to allow their smartphones or other mobile devices to be used to transport such advertising content. One factor in this dearth of useful data is the general difficulty in capturing data from smartphone users. New measurement studies have recently been conducted (Do and Perez, 2010; Falaki et al, 2010; Shye et al, 2009) and testbeds built (Shepard et al, 2010), but these tend to concentrate on network-level data such as traffic statistics, rather than user-level data such as willingness to participate in content sharing.

A second factor is that collecting data from advertising networks is fraught with difficulties itself. Guha et al.(Guha et al, 2010) outline the challenges in measurement, including differences between measurement clients, due to noise from DNS load-balancing, and the churn of advertisements. These difficulties become even more challenging in a mobile environment: a user study that proposed to measure and collect data from mobile users in a mobile advertising environment would no

---

[2] `http://advertising.apple.com/`

doubt have even more variations in advertising, given differences between users in locations, activities, behaviours and networks.

One solution might be to combine passive measurement with an active user study, allowing experimental participants to verify the types of advertisements being received, and the contexts, behaviours and experiences under which they are willing and unwilling to view, click on, or distribute such advertisements. One mechanism for doing this might be to use the Experience Sampling Method, where participants combine a diary study with signal-contingent alerts which trigger questions (Larson and Csikszentmihalyi, 1983). For instance, a participant could be detected near a shopping centre or restaurant, and this might trigger questions about their advertising preferences. Experience Sampling has been combined successfully with mobile devices, e.g., (Mancini et al, 2009; Consolvo and Walker, 2003; Ben Abdesslem et al, 2010), but care should be taken before implementing a large-scale mobile advertising ESM study. For instance, the ethical, legal and privacy implications of collecting mobile location and advertising data need to be considered (Henderson and Ben Abdesslem, 2009), and mechanisms for anonymising data accordingly need to be designed, especially given the number of related datasets that have been subsequently deanonymised by researchers (Ohm, 2009).

There are a number of solutions for avoiding click fraud and performing better advertisement. One suggestion is to charge based on user's actions, i.e., the publisher gets a premium only after the successful conversion of the advertisement, meaning the user's visit to the advertiser's website and performing an action such as buying an item or signing up for a service. There are a number of basic attempts at such an approach by means of tracking cookies, however these efforts make up a negligible portion of the current advertising revenue on the Internet.

Juels *et al.* (Juels et al, 2007) propose a cryptographic approach for replacing the pay-per-click model with one in which pay-per-action (e.g., shopping) can attract premium rates and unsuccessful clicks are discarded. in this system, the users which make a purchase are identified by the network of advertisers as premium advertisers. The client browsers use a coupon instantiated by third party cookies or issued by the attestor upon redirection. The disadvantage of this method is the ability of malicious attacker, possibly an advertiser, to use a botnet and replay the coupons numerous times, for a large number of cooperating publishers. This will then force the syndicator either discount all those replays, or removing those clients from the system with valid coupons. In both cases, the advertisement income is minimised. It also allows for the syndicator and the attestor (ad broker and middle box) to profile the users accurately including their spending budget. They also indicate that most standard click fraud techniques remain unsolved today. Despite early suggestions of this method, it has not been implemented on a large scale as it requires trust between advertisers and publishers.

Some advertisers have suggested the use of anonymised ISP data streams for verification of clicks and for better user profiling. Attempts to do so, such as Phorm, have been unsuccessful due to user privacy concerns. Privacy reasons also prevent brokers from releasing their server logs and click data to advertisers and their agents for deep inspection of the click rates. Other solutions include use of human-invisible

advertisements to act as traps for botnets, but these can easily be ignored by a simple visibility test in botnet design.

Immorlica *et al.*analyse the click-fraud learning algorithms to compute the estimated click-through rate (Immorlica et al, 2005). They focus on a situation in which there is just one advertisement slot, and show that fraudulent clicks can not increase the expected payment per impression by more than $o(1)$ in a click-based algorithm. The complexity of the inferred algorithm and the need for click-through rate estimation, however, would make it impractical as it also deviates from the pay-per-click model, to pay-per-view model, which is the least desired model in the modern advertisement world where bidding for space is of critical importance.
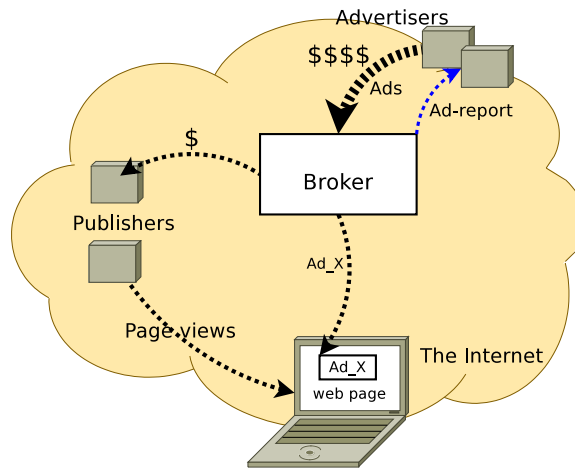
3 Internet Targeted Advertising Basics

Fig. 1 Targeted keyword-based advertising.

Before introducing our MobiAd system, we first characterise today's advertising systems. These can be broken down into four major components: *advertisers*, *publishers*, *clients*, and *brokers*.

Advertisers wish to sell their products or services through advertisements. Publishers (e.g., news and review websites, personal weblogs) provide opportunities to view advertisements, for instance by providing space for advertising banners. Clients are the devices that show publisher web pages and advertisements to users. Brokers (e.g., Google or Yahoo!) bring together advertisers, publishers, and clients. They provide advertisements to users, gather statistics about which advertisements were shown on which publisher's pages, collect money from the advertisers, and pay the publishers.

Figure 1 illustrates the most popular advertising model on the Internet today. Advertisers specify their advertisements and bids (how much the advertiser is willing to pay for views and clicks of the ads) to the broker. When a publisher provides banner space to the client on a web page, a request goes to the broker, asking it to fill in the banner space with appropriate advertisements. The provider makes the decision as to which advertisements to place based on a number of criteria such as the keywords for the web page, personalisation information about the client (usually persistent cookies on client machine), the keywords of the advertisement, and the bid associated with the advertisement. It then delivers the advertisement to the client, informs the advertiser of the advertisement's views and clicks, and charges the advertisers and compensates the publishers accordingly.
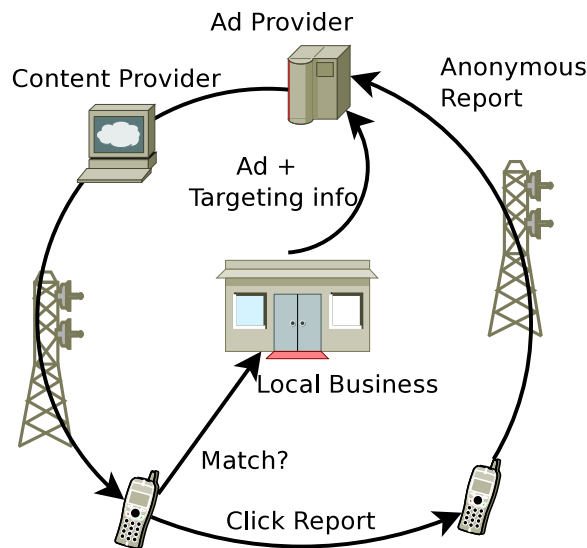
4 MobiAd architecture

Fig. 2 Targeted localised advertising.

In this section we present an overview of the MobiAd system architecture. Figure 2 presents an example of some of the components of the MobiAd advertising system. The main components of the system are:

- *Advertisers*: The advertising providers aim to reach a specific groups of users based on requirements such as sex, age, interests and location. They provide

the advertisement texts, targeting information, bidding budgets and potentially a landing page for the clicks[3].

- *Network Operator*: The network operator provides the infrastructure for disseminating the advertisements, collecting reports and locating users. In return they receive a share of revenues.
- *Content Provider*: Content providers, such as news sites and blogs, provide online services and materials which are of interest to users. Alongside their main content, they provide advertising boxes where personalised advertisements can be displayed to the user. Social networking sites such as Facebook are a particular type of publisher, with access to detailed profile information on their users that can be used for targeting.
- *Ad Provider*: The ad providers (e.g., Google or Microsoft) are the interface between the advertisers, operator and content providers. They gather advertisements from advertisers, provide advertisements for the users in the publisher websites, collect view and click reports, bill the advertisers and compensate the publishers and the network operator.
- *Profiling Agent*: The agent gathers relevant information for profiling users. It also downloads and filters relevant advertisements, displays them to the user at appropriate and convenient times and prepares click and view reports for billing purposes.

By using detailed profiling and data-mining techniques in addition to sensed mobile information such as users' locations, MobiAd provides new opportunities for localised and personalised advertising. In terms of privacy protections, some of these components are also similar to traditional advertising systems (such as Google's AdWords program) or newly proposed privacy-aware systems such as Adnostic (Toubiana et al, 2010) and Privad (Guha et al, 2009). The key difference in our system, however, here is the fact that mobility and use of local advertisement distribution and Delay Tolerant Networks (DTN) (Fall, 2003) provide a simple and scalable advertisement distribution and privacy-preserving click-report mechanism, while addressing many of the numerous challenges for profiling and advertisement placement on a mobile device with limited screen size and battery life. The network operator also plays a more central role as it needs to broadcast advertisements in a localised manner and collect and forward reports. The lower bandwidth, battery life and display size of mobile phones prevent us from downloading, sorting and showing a large number of advertisements on the user's phone. Hence in MobiAd we focus on displaying a lower number of advertisements, but with higher targeting and a focus on local advertisements that would particularly benefit from the user's location information.

In the next sections we expand on the key individual components, their roles and operation strategies.

---

[3] We elaborate in Section 7 why a landing page is not essential

5 Profiling and Incentives

The most important objective of MobiAd is to serve relevant and interesting advertisements to the user. Since the mobile phone's battery life and display size and general browsing time are currently reduced compared to the average personal computer, it is crucial to use the advertisement display opportunities effectively. In order to do this, users' interests and profiles should be maintained on user handsets, while allowing the user to configure and delete their interest categories. This is also in compliance with requirements and recommendations of most regulatory organisations and privacy advocates.

5.1 Maintaining the User Profile

There are many rich sources of profile information on a typical smartphone, from email and browsing activities to social networking and shopping sites. This information is in essence an aggregation of information from the user's web history, application caches and keyword extractions from activities on social networks and email. Users are likely to have different privacy sensitivities regarding these various data sources, and should be allowed to control which are included in profiling activities. Browsing behaviour can be used to update profiles at lower processing cost using server-side pre-categorisation of URIs into interest segments (Toubiana et al, 2010).

The profile and the associated software work in cooperation in a similar manner to Google's Gmail or persistent cookies from search engines and advertising providers. In MobiAd, however, the profile does not leave the user's handset and the software platform picks up the appropriate advertisements from the broadcast channel.

As a further protection the profile is an aggregate view of user interests rather than a detailed history, reducing the risk of information leakage. Such aggregated information can exclude information about sensitive matters such as medical interests, trade union membership and religious beliefs. This builds user trust in the system, reduces the potential for this information to be accessed for unauthorised purposes, and enables easier compliance with data protection laws such as the European Union's Data Protection Directive (European Parliament, 1995).

In this design, user profiles are kept solely and securely on the handset. The profile must be visible to the user but unobtainable by other applications. The isolation of information between different applications is readily available on popular smart phones. Profiling tasks can be done while the phone is idle. The extent and depth of categorisation is dependant on the different regions and users, e.g., Google keeps 700 categories in a 3-level hierarchy, while Amazon has over 65,000 categories. We envisage that a MobiAd client can maintain an extensive database of interests, locations, mobility patterns and daily habits. Such detailed information would enable the relevant advertisements to be easily filtered and directed to the user.

5.2 User Incentives

The MobiAd system is clearly beneficial to advertisers and network operators, but why would users install such an application? Users have an incentive to install and utilise most applications if there is a marginal entertainment or financial benefit for them. iPod touch users download an average of 12 apps a month and spend 100 minutes a day using apps.[4] Android and iPhone users download a similar number of apps every month and spend a similar amount of time using the apps (AdMob, 2010). On a new iPhone app, users have been reported to be searching daily for money saving vouchers and local promotions.[5] Hence the intention is that useful services would encourage the users to download the client which could also act as a privacy information centre on their phone.

For the MobiAd system, we are exploring a range of advertisement benefits to the user, location-based and independent. Location-based benefits could include offers and coupons for local businesses and retailers. Independent long-term benefits could include informative applications, such as suggesting events and activities which could be off interest to the user and are not necessarily advertised. In addition, network operators may pre-install this type of software on handsets, or offer incentives to users (such as discounted monthly fees) for them to use MobiAd. In this way the costs of carrying other user reports can also be compensated by the *availability* of a user's handset for carrying traffic and hence contributing to the anonymisation process. Another incentive could be a small percentage cut payment from the advertising click revenue for the report carriers, in an aggregate manner, so as to avoid the network operator or the advertiser being able to trace the origin of the clicks.

6 Dissemination and Reporting

6.1 Advertisement Dissemination

Dissemination of advertisements in a mobile environment is different from the desktop environment. In MobiAd, the focus is on local advertisements that are relevant to the user. Location information can be obtained using GPS, Wi-Fi or network provider information from the handset. While users may roam in and out of mobile cells and thus affect advertisement download rates, it has also been shown that there are limits to predictability of location of users at given times (Song et al, 2010). We therefore do not rely heavily on prefetching all the relevant advertisements to the user, apart from at locations such as home and work where they appear frequently.

The optimal data dissemination strategy should avoid constant data download, but be ready for unpredictable arrival of the user into new areas. The MobiAd agent

---

[4] http://www.appleinsider.com/articles/10/02/25/ipod_touch_users_
spend_more_time_using_apps_than_those_with_iphones.html
[5] http://www.pocket-lint.com/news/34077/deals-moneysupermarket-launches-iphone-app

on-handset should be able to classify locations that are frequently visited (using a list of GPS positions most frequently visited), such as the route from home to work, weekend hotspots and such like. The advertisements for these locations could be prefetched when there is wireless connectivity and stored for longer periods, to minimise data transfer costs and battery utilisation on the handset.

When a user enters a new location where the relevant advertisements have not already been prefetched, she can receive all local advertisements using technologies such as Multimedia Broadcast and Multicast Services (MBMS) (MBMS, 2010). MBMS is a new service offered on GSM and UMTS networks and uses multicast distribution in the core network that enables an interaction between the handset and the network which can be used for distributing advertisements and collection of reports. MBMS enables network operators to distribute all the local advertisement texts simultaneously to all cell phone users within each transmitter's coverage area using a single shared transmission broadcast. As cell coverage is expected to be in order of a few hundred metres, the text advertisements within each cell should not exceed a few kilobytes (a few hundred local advertisements, each having around 100 characters of text) which is a reasonable amount of data transfer for all modern smart phones to deal with. If one channel is used at each cell tower to broadcast locally-relevant ads, all phones could listen to all channels and just select relevant advertisements without revealing which advertisements are of interest and shown to the user. If more privacy is required, protocols such as SlyFi (Greenstein et al, 2008) could be used to provide anonymous sniffing capabilities for downloading advertisements from local Wi-Fi hotspots. However since all the local advertisements can be downloaded from Wi-Fi hotspots, the hotspot service provider is not able to classify the users, as they cannot find out which advertisements were displayed. The number of broadcast advertisements even in busy metropolitan areas could be limited by a combination of network operator and ad provider using an auction mechanism to limit the number of location-targeted advertisements through raising their price. Hence we opted for local flooding for the current design.

It is also possible for the advertising network to suggest a shortlist of relevant advertisements within specific advertisement frames, resulting from centrally-available context and pricing information (Toubiana et al, 2010). One may also consider using Tor[6] for ad dissemination or report collection, however Tor requires a real-time interactive channel that consumes significant amounts of power. MobiAd's transmission of advertisement reports does not need a real-time interactive anonymous channel. By relaxing this requirement, we can reduce power use.

---

[6] http://www.torproject.org/

6.2 Billing

At the end of each billing cycle, advertisers are billed by the advertising network for advert displays and click-throughs. MobiAd uses a cryptographic protocol developed by Toubiana et al. that allows clients to notify the network of advertisement impressions without leaking user interest information (Toubiana et al, 2010).
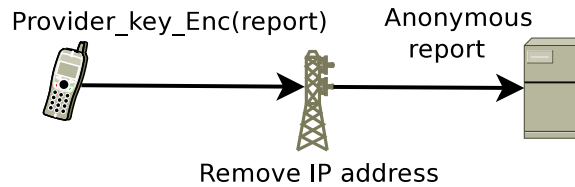

Fig. 3 Encrypting click reports.

Figure 3 provides an overview of the public key encryption stages of the click reports. The advertisement report will be encrypted using the advertising provider's public key, so only the advertising provider can open the report. Advertisement clicks will then be anonymised, as the advertising provider can identify the advertisement which was clicked, but does not know who clicked on the advertisement. Likewise, the network operator knows a report was received, but does not know what advertisement was clicked on. MobiAd also uses a one-time pseudorandom number in order to avoid replay clicks. This is similar to the reporting mechanism used in Privad (Guha et al, 2009). We avoid using more sophisticated methods such as Tor due to the complexity of running such CPU- and data-intensive systems on mobile phones.

6.3 Report Collection Using DTN

As advertisers are generally billed using information on cost-per-impression or cost-per-click, there needs to be a return path for clients to report these data (without leaking information on user interests). To further protect users against attempts to link reports to user behaviour, we take a similar approach to onion routing (Dingledine et al, 2004) using the DTN paradigm.

DTN was originally designed for interplanetary communication, where the delays can be several minutes or longer (Burleigh et al, 2003), and then was adapted to solve intermittent connectivity problems in daily life. Furthermore, it has been recently shown that by leveraging the delay of transmission, DTN can improve the anonymity of wireless communication from physical localisation (e.g., triangulation) (Lu et al, 2010). Onion routing is an approach to achieve anonymous communication by routing messages through several intermediate relays before reaching

their destinations, and so the probability of revealing the source node of a message is significantly reduced. As shown in Figure 4, the MobiAd agent system is designed to report on advertisement views and clicks, while preserving the privacy and anonymity of users.
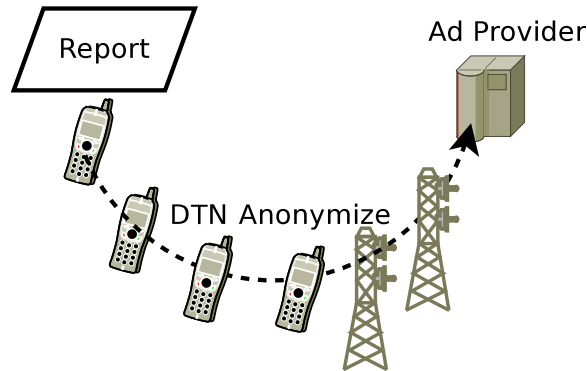

Fig. 4 Collecting reports via a Delay-Tolerant Network.

In MobiAd, we use DTN to route advertisement reports to several intermediate relays before they are finally passed to the cellular network. DTN relies on mobile peer-to-peer store-and-forward (using Wi-Fi or Bluetooth connections) and hence there is no additional monetary cost on top of the cellular network cost. Here for further privacy consideration we have three requirements: 1) the relays should have certain social in-correlation with the social network, which prevents identity reverse engineering from the social relationship, 2) if possible, the final location of the final hop to the cellular network should have certain geographical distance from the original location of the report, 3) we want a certain delay (but not too long for billing purpose) between the time when the report is first sent out from the source and the time when it is finally sent to the network.

To better guarantee successful delivery, we use two-copy forwarding instead of a single copy (Spyropoulos et al, 2004) for the reports, which means that the source will make a duplicate copy of the report and send them separately to two different relays. The report will not be further duplicated during the multiple-hop transmission. To achieve the social in-correlation criteria, we take an anti-social network approach as opposite to the social-based approach introduced in BUBBLE Rap forwarding (Hui et al, 2008), where a mobile device will periodically scan the environment and detect the devices belonging to its social community. But here we obfuscate the sender's social network by adding random nodes, including friends and strangers, as relays instead of socially-close nodes (Parris and Henderson, 2010). In this way the social network of a node cannot be easily identified by an attacker who monitors all the forwarded packets in a cell network where a node is frequently present.

In order to preserve the location and temporal privacy, we will set the number of hops before the last hop to the cellular network to be 3 (so in total 4 hops). Based

on the seminal work on 6-degree of separation by Milgram et al. (Milgram, 1967), 4 hops should be a reasonable distance in order to scramble the social correction, and long enough for the message to have enough temporal delay and geographical distance from the source. There may be energy consumption issues due to excessive wireless scans for efficient DTN routing, but since delay is not a main concern for the delivery of the advertisement reports, we do not need to scan the environment so frequently.

While taking all these measures into consideration, it is theoretically possible for an advertiser or the network operator, through long-term monitoring of the mobile user, to determine which advertisement reports have a geographical correlation with the user's location. This is due to the user's routes following a specific home-work-home pattern for most of the days. In order to overcome this, MobiAd can employ a system where base stations follow a similar approach to the DTN system proposed before forwarding the reports for, for example reports from a specific region or town could be forwarded all over the country, or they could be presented to the advertising provider in aggregate form. In this way the advertising provider cannot build an accurate estimate of number of phones and their geographic correlation in specific regions. We are currently working towards categories of attack scenarios by the advertisers, advertising provider, network operator and content provider and plan to address these issues in future work.

7 Security and Privacy Challenges

MobiAd is privacy-aware, but several security and privacy challenges remain. An open question is whether users that click on advertisements should be taken directly to an advertiser's URI; redirected to an advertiser site via an intermediate URI hosted by the ad provider for click-through measurement and fraud protection, as commonly happens in today's advertising networks; or to content also distributed using an anonymous channel, to further limit the potential for linking users to specific interests. In general usage the click-through rate for adverts is extremely low, so sending users directly to advertisers' sites is much less privacy-intrusive than building detailed server-side user profiles. Users may anyway voluntarily provide further information to advertisers at this point, particularly if they make a purchase. However, particularly privacy-sensitive users may make use of a service such as Tor to reduce linkage of their browsing behaviour to any long-term identifier (such as an IP address). Care needs to be taken to reduce the ability of malicious advertisers to gain information on users who click-through an advert that is targeted at extremely small numbers of individuals – both in terms of interests and in frequently-visited locations such as homes and work places.

Careful attention also needs to be paid to client-side implementation details to prevent information leakage. Advertisements need to be carefully isolated within display pages using mechanisms such as identically-sized iframes, to prevent client-

mediated communications between publisher and advertiser. This may preclude the inclusion of active content such as Flash ads (Toubiana et al, 2010).

Mobile handsets are less frequently shared than PCs, and hence information is less likely to leak between users of the same equipment. However, care must be taken to protect profiles using a PIN or password from access by other people with physical access to a handset. The possibility of coerced access – such as by parents to their children's handset profile – must also be considered, which is a further reason for storing only aggregate information and excluding sensitive personal data categories.

An issue outside the scope of MobiAd is users' reactions to highly-targeted adverts, even with guarantees that behavioural profiles remain entirely private to the individuals they describe. Advertisers may need to tread carefully in targeting adverts for products such as low-fat foods that to some users may raise concerns that they have been unfairly categorised, or suggest lifestyle problems. A possible mechanism to address the first concern would be transparency in explaining to users why they had been shown any given advert. Many countries also have laws that ban discriminatory treatment of individuals based on certain characteristics that might be inferred from behavioural profiles. For sensitive advertisements we envisage that no reports need to be collected in order to minimise any privacy leaks. Even the landing pages of such advertisements could be provided using a Content Distribution Network (CDN), or they can be pre-fetched using the DTN system. We are dealing with the privacy issues in more detail for future work.

8 Detecting Click Fraud

One of the main reasons that online advertising platforms keep detailed logs of user clicks on adverts is to detect instances of fraud. In this section of this chapter we describe some of the mechanisms used in such click-through fraud, and techniques such as Bluff Ads that can be used to ensure the extra privacy protections of MobiAd are not abused to defraud advertisers.

For fraud detection, one can add some untargeted ads, or as we refer to them, Bluff ads to the advertisements displayed to the user. These are real advertisements, but served randomly. Every time the user visits a publisher page, we serve the user with probability $p(i)$ with profiled advertisements, and with probability $[1 - p(i)]$ with other, non related Bluff ads. The brokers' entire advertising model is based on the idea of showing only the most relevant advertising content. If displayed advertisements are poisoned with context-free advertisements on a frequent basis, benign users will perceive this as the broker doing a poor job at finding relevant advertisements. Hence, the Bluff/real ratio must be set in a way that the user's browsing experience and advertising quality perception is not greatly affected. For example, a user living in Iran should not ideally be presented with special offer advertisements on beer during Oktoberfest. But it might not unreasonable to be shown car adverts, though his profile has no indication of his interests in driving. In practice,

the Bluff ads should be authentic advertisements of different advertisers, spread in the network and shown randomly, but never charged for.

The Bluff ads serve two purposes. First, they give the user "comfort" that he is not being watched too closely and monitored too deeply. Secondly and more importantly, they help identify fraud clickers and eliminate them from the system. These fraud agents are just clicking for publishers, or against a specific advertiser.

8.1 Using Bluff Ads

We now address different forms of click-fraud attacks and those on user privacy, and we briefly describe how the Bluff ads will help minimise them.

8.1.1 Profiling the customer

The Bluff ads will prevent publishers and advertisers from narrowly targeting the client as it is not possible to know whether the viewed advertisement was a Bluff or not. Unless the publishers work in large groups together which will also increase the difficulty level for them. The advertisers can naturally profile users easier, however that is not avoidable as ultimately users; interests lead to revenue generation for the advertiser and broker. If the broker notices that a specific web page covers many categories it can ignore that website altogether. In the case of content aggregators the broker can put specific emphasis on pages visited during client;s browsing session.

8.1.2 Publisher fraud

The most common case is where a publisher has hired a large botnet to perform clicks for it. This can be easily realised from the frequency of the clicks. In this case the publisher ID and the Bluff/real ratio for publishers can inform the broker of this attack. If the Bluff/real ratio is higher than an average user, it is indicative of a bot being in operation.

8.1.3 Attacks on advertisers

These can again be identified by the a combination of simple threshold sampling and Bluff/real ratio of the advertisements. If most of the clicks from a host are targeted towards a single advertiser, there will be an obvious trend in their Bluff/real ratio. If the attacker decides to poison the statistics by clicking on random other advertisements, the ratio will be affected again. Many large advertisers today use specialised agencies to monitor their incoming traffic and identify spammers and click-fraud users, who tend to visit often but spend no time on the advertiser website.

These users are also identifiable if they come from same IPs with frequent visit counts (simple threshold detection). The advertisers' agents can hence pass a list of fraud suspects to the broker who will remove them from the billing system.

8.1.4 Attacks on publisher

These are the most difficult attacks, when a publisher is under attack from another source. Such attacks happen when reliable publishers, such as CNN, who use advertising brokers, are under attack in order to damage their relationships with the advertisers and providers and ultimately eject them from the competition scene when bidding for advertising space on the page. It is possible to detect such attacks by examining the Bluff/real ratio for the advertiser and publisher pair and identify these if the frequency of views/clicks is less than a certain time threshold. Distributed attack on publishers are a new form of attack and further research is needed to determine solutions for detecting these in detail.

9 Summary

In this chapter we presented the MobiAd system architecture, a system for delivering personalised, localised and private yet scalable mobile advertisements. In this system, advertisements are locally broadcast to users within mobile cells, appropriate advertisements are shown to the user and view and click reports are collected using a DTN system, preserving the privacy and anonymity of the user. MobiAd provides an opportunity for using the significant amount of information on users' smart phones for targeted advertising while protecting their privacy.

We also presented a brief overview of the current challenges in detection of click-fraud in online advertising. We presented a simple detection strategy, Bluff ads. These are sets of irrelevant advertisements displayed amongst user's targeted advertisements, which should not be clicked on by an authentic user. Together with threshold detection, IP address monitoring and profile matching techniques, Bluff ads can be used to make it raise the bar for botnet owners to train their software, or a human operator. The Bluff ads also may have a comfort factor of decreasing the user's negative perceptions by reducing the number of accurately targeted advertisements. We are currently working on deployment of Bluff advertisements on a large advertising service.

References

Aalto L, Göthlin N, Korhonen J, Ojala T (2004) Bluetooth and WAP push based location-aware mobile advertising system. In: Proceedings of the 2nd international conference on Mobile

systems, applications, and services (MobiSys), ACM, New York, NY, USA, pp 49–58, DOI 10.1145/990064.990073

AdMob (2010) Admob mobile metrics report, `http://metrics.admob.com/wp-content/uploads/2010/02/AdMob-Mobile-Metrics-Jan-10.pdf`

Ben Abdesslem F, Parris I, Henderson T (2010) Mobile experience sampling: Reaching the parts of Facebook other methods cannot reach. In: Proceedings of the Privacy and Usability Methods Pow-Wow (PUMP), British Computer Society, URL `http://scone.cs.st-andrews.ac.uk/pump2010/papers/benabdesslem.pdf`

Burleigh S, Hooke A, Torgerson L, Fall K, Cerf V, Durst B, Scott K, Weiss H (2003) Delay-tolerant networking: an approach to interplanetary internet. IEEE Communications Magazine 41(6):128–136, DOI 10.1109/MCOM.2003.1204759

Consolvo S, Walker M (2003) Using the experience sampling method to evaluate ubicomp applications. IEEE Pervasive Computing 2(2):24–31, DOI 10.1109/MPRV.2003.1203750

Cozza R, Nguyen TH, Gupta A, Vergne HJDL, Sato A (2008) Market Share: Smartphones, World-wide, 4Q08 and 2008

Dingledine R, Mathewson N, Syverson P (2004) Tor: The second-generation onion router. In: In Proceedings of the 13th USENIX Security Symposium, pp 303–320

Do TM, Perez DG (2010) By their apps you shall understand them: mining large-scale patterns of mobile phone usage. In: Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia (MUM), ACM, New York, NY, USA, DOI 10.1145/1899475.1899502

European Parliament (1995) Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. OJ L 281 pp.31-50

Falaki H, Lymberopoulos D, Mahajan R, Kandula S, Estrin D (2010) A first look at traffic on smartphones. In: Proceedings of the 10th annual conference on Internet measurement (IMC), ACM, New York, NY, USA, pp 281–287, DOI 10.1145/1879141.1879176

Fall K (2003) A delay-tolerant network architecture for challenged internets. In: SIGCOMM '03: Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications, ACM, New York, NY, USA, pp 27–34, DOI 10.1145/863955.863960

Freudiger J, Vratonjic N, , Hubaux JP (2009) Towards Privacy-Friendly Online Advertising. In: Proceedings of W2SP 2009: Web 2.0 Security and Privacy, URL `http://w2spconf.com/2009/papers/s2p1.pdf`

Greenstein B, McCoy D, Pang J, Kohno T, Seshan S, Wetherall D (2008) Improving wireless privacy with an identifier-free link layer protocol. In: MobiSys '08: Proceeding of the 6th international conference on Mobile systems, applications, and services, ACM, New York, NY, USA, pp 40–53, DOI 10.1145/1378600.1378607

Guha S, Reznichenko A, Tang K, Haddadi H, Francis P (2009) Serving Ads from localhost for Performance, Privacy, and Profit. In: HotNets-VIII: Proceedings of the Eighth ACM Workshop on Hot Topics in Networks, URL `http://conferences.sigcomm.org/hotnets/2009/papers/hotnets2009-final27.pdf`

Guha S, Cheng B, Francis P (2010) Challenges in measuring online advertising systems. In: Proceedings of the 10th annual conference on Internet measurement (IMC), ACM, New York, NY, USA, pp 81–87, DOI 10.1145/1879141.1879152

Haddadi H (2010) Fighting online click-fraud using bluff ads. ACM SIGCOMM Computer Communication Review 40(2):21–25, DOI 10.1145/1764873.1764877

Henderson T, Ben Abdesslem F (2009) Scaling measurement experiments to planet-scale: ethical, regulatory and cultural considerations. In: HotPlanet '09: Proceedings of the 1st ACM International Workshop on Hot Topics of Planet-Scale Mobility Measurements, ACM Press, New York, NY, USA, pp 1–5, DOI 10.1145/1651428.1651436

Hui P, Crowcroft J, Yoneki E (2008) BUBBLE rap: Social-based forwarding in delay tolerant networks. In: MobiHoc '08: Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing, ACM, New York, NY, USA, pp 241–250, DOI 10.1145/1374618.1374652

Immorlica N, Jain K, Mahdian M, Talwar K (2005) Click Fraud Resistant Methods for Learning Click-Through Rates. In: Deng X, Ye Y (eds) Internet and Network Economics, Springer Berlin / Heidelberg, Berlin, Heidelberg, Lecture Notes in Computer Science, vol 3828, pp 34–45, DOI 10.1007/11600930_5

Juels A, Stamm S, Jakobsson M (2007) Combating click fraud via premium clicks. In: SS'07: Proceedings of 16th USENIX Security Symposium on USENIX Security Symposium, USENIX Association, Berkeley, CA, USA, pp 1–10

Komulainen H, Ristola A, Still J (2006) Mobile advertising in the eyes of retailers and consumers - empirical evidence from a real-life experiment. In: Proceedings of the International Conference on Mobile Business, IEEE Computer Society, Washington, DC, USA, pp 37–, DOI 10.1109/ICMB.2006.31

Larson R, Csikszentmihalyi M (1983) The experience sampling method. New Directions for Methodology of Social and Behavioral Science 15:41–56

Lu X, Hui P, Towsley D, Pu J, Xiong Z (2010) Anti-localization anonymous routing for Delay Tolerant Network. Computer Networks 54(11):1899–1910, DOI 10.1016/j.comnet.2010.03.002

Mancini C, Thomas K, Rogers Y, Price BA, Jedrzejczyk L, Bandara AK, Joinson AN, Nuseibeh B (2009) From spaces to places: emerging contexts in mobile privacy. In: Ubicomp '09: Proceedings of the 11th international conference on Ubiquitous computing, ACM, New York, NY, USA, pp 1–10, DOI 10.1145/1620545.1620547

MBMS (2010) Multimedia Broadcast/Multicast Service (MBMS); Stage 1, 3GPP Specification. URL http://www.3gpp.org/ftp/Specs/html-info/22146.htm

Merisavo M, Vesanen J, Arponen A, Kajalo S, Raulas M (2006) The effectiveness of targeted mobile advertising in selling mobile services: an empirical study. International Journal of Mobile Communications 4(2):119–127, URL http://www.metapress.com/content/4RE0HR5YAARJC061

Milgram S (1967) The small-world problem. Psychology Today 1(1):61–67

Ohm P (2009) Broken promises of privacy: Responding to the surprising failure of anonymization. Social Science Research Network Working Paper Series URL http://ssrn.com/abstract=1450006

Parris I, Henderson T (2010) Privacy-enhanced social-network routing. Computer Communications DOI 10.1016/j.comcom.2010.11.003

Ranganathan A, Campbell RH (2002) Advertising in a pervasive computing environment. In: WMC '02: Proceedings of the 2nd international workshop on Mobile commerce, ACM, New York, NY, USA, pp 10–14, DOI 10.1145/570705.570708

Shepard C, Tossel C, Rahmati A, Zhong L, Kortum P (2010) Livelab: Measuring wireless networks and smartphone users in the field. In: Proceedings of The 3rd Workshop on Hot Topics in Measurement & Modeling of Computer Systems (HotMetrics), URL http://hotmetrics.cs.caltech.edu/pdfs/paper12_final.pdf

Shye A, Scholbrock B, Memik G (2009) Into the wild: studying real user activity patterns to guide power optimizations for mobile architectures. In: Proceedings of the 42nd Annual IEEE/ACM International Symposium on Microarchitecture, pp 168–178, DOI 10.1145/1669112.1669135

Song C, Qu Z, Blumm N, Barabasi AL (2010) Limits of Predictability in Human Mobility. Science 327(5968):1018–1021, DOI 10.1126/science.1177170

Spyropoulos T, Psounis K, Raghavendra C (2004) Single-copy routing in intermittently connected mobile networks. In: Proceedings of the First Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON), IEEE, pp 235–244, DOI 10.1109/SAHCN.2004.1381922

Toubiana V, Narayanan A, Boneh D, Nissenbaum H, Barocas S (2010) Adnostic: Privacy preserving targeted advertising. In: Proceedings of the 17th Annual Network and Distributed System Symposium, Internet Society, San Diego, California, USA

Turow J, Hennessy M (2007) Internet privacy and institutional trust. New Media & Society 9(2):300–318, DOI 10.1177/1461444807072219