Privacy-Preserving Targeted Mobile Advertising Using Federated Identity Management Systems

Waleed A. Alrodhan

College of Computer and Information Sciences Al-Imam Muhammad ibn Saud University

Summary
Mobile advertising is taking over the business of advertising after a rather rapid leap in both revenue and efficiency. Targeted mobile advertising (TMA) is one of the most efficient models of mobile advertising; however, it raises the issue of privacy preserving and data protection. In order to address that, a scheme named Privacy-Preserving TMA (or PPTMA) has been recently proposed that included practical solutions to many privacy issues associated with TMA. However, we believe that PPTMA can be more practical and secure if combined with a Federated Identities Identity Management (FIdM) system. In this paper, we prove that this proposed integration would minimize the computational and storage requirements and boost up the security level. In FIdM, the identity provider is trusted by definition, and security tools and techniques like for example pseudonyms, cryptography, secure messaging, strong authentication, and many more, are already embedded. Moreover, integrating PPTMA with widely used FIdM systems would increase its practicality and the user acceptance.

Key words:
Privacy, security, mobile, advertisement, identity.

1. Introduction

Digital advertising is rapidly growing. Total spending on digital advertising has jumped from $19.20 billion in 2013 to reach $68.69 billion in 2015, and expected to reach $195.55 billion in 2019 [1]. In the US, it is highly expected that the spending on digital advertising will exceed the spending on TV advertising in 2019 [2].

In the realm of digital advertising, mobile advertising has made giant growth in the past few years; according to many studies, mobile advertising is currently leading the business of digital advertising. For example, a study in [3] shows that mobile advertising has overtaken desktop advertising in 2013. Another study in [4] concluded with very similar results. If we take into consideration that the average user spends 90% of her "mobile usage time" on mobile apps and only 10% on the browser [5], we can readily deduce that mobile advertisers spot their focus on mobile apps as their advertising courier.

In the mobile advertising ecosystem, an advertiser pays an ad-network to publish its advertisement. The ad-network shares the payment with app-developers who embed the advertisement in their apps. Many app-developers provide their apps for free and rely on mobile advertising for revenue [6]. A study published in [7] shows that 46 out of 60 off-the-shelf apps selected randomly from the top free apps in Google Play Store had at least one ad-library embedded in it.

In this paper we propose an advertisement system that would deliver tailored mobile advertisements to the users along with preserving their privacy via a practical and efficient model. This system is based on an integration of a federated identities management system with a privacy-preserving targeted mobile advertising system.

The remainder of this paper is organized as follows. Section 2 provides an overview of targeted mobile advertising, privacy-preserving targeted mobile advertising, and federated identities management. In Section 3 we propose our mobile advertising model and in Section 4 we provide an analysis of it. In Section 5 we describe a prototype implementation of the proposed model. Section 7 discusses possible future work. Finally, section 7 concludes the paper.

2. Overview

In this section we provide a brief overview of the mobile advertisement process within the targeted mobile advertising (or TMA) and the privacy preserving targeted mobile advertising (or PPTMA) systems. Also we generally discuss the concept of federated identities management (or FIdM) system.

2.1 Targeted Mobile Advertising

The main concept of TMA is to automate the mobile advertisement selection process based on the targeted user's interests. Whilst this raises many privacy violation concerns, it boosts the efficiency of the whole advertisement system by showing the most suitable and relevant advertisements to its users. Also, this means more revenue for both the ad-network and the developer, along with a better RoI for the advertiser. Figure 1 shows the framework of the TMA. Basically, each ad-network owns its own TMA system in which the user's information (e.g. personal profile and interests) are collected by specific plug-ins embedded in the participant apps. This information is then sent to the ad-network so it can pick the most relevant advertisement up from the ad-pool it maintains.
Because of the privacy violation concerns associated with the TMA systems, many mobile users end up blocking their ads. In order to address this issue, PPTMA was proposed in [7]. PPTMA preserves users' privacy and keeps the benefits of the TMA at the same time.

2.2 Privacy-Preserving Targeted Mobile Advertising

It needless to mentioned that in order to come up with the most suitable ad to the user, a TMA needs to first collect personal information about her, like for example, her age, gender, location, visit-history, etc. Such information is classified 'personal' in the OECD guidelines, and should be protected [8]. This raises many security and privacy concerns and will mostly result in the user faking her information or the operating system blocking all ad's [9, 10]. However, many studies showed that there are many users' who welcome such ads in case they get rewarded for it (e.g. free app) [11, 12, 13]. PPTMA was proposed to solve this dilemma: keeping targeted advertising whilst protecting users' personal information [7]. Whilst PPTMA, as TMA, collects personal information to come up with the most relevant ad, it ensures that users' private information is not shared with any party in the echo-system (e.g. the app-developer). It remains in the users' mobile phones, and can only be shared if there is an explicit user consent to do so. Moreover, users' have the ability to opt-in or opt-out of the system at any time.

PPTMA can be seen as a local component that must be installed on the user's mobile phone. It includes four main modules:

- **User profile manager**: responsible of collecting user's personal information based on the privacy policy.
- **Data access manager**: responsible of managing the access to the collected personal information.
- **Ad-plugins scanner**: responsible of scanning and third-party plugins for malicious code of behaviour.
- **Ad-selector**: responsible of selecting the most relevant ad from the ad-pool, and of generating trusted billing records.

Briefly, PPTMA work as follows:

1. The user profile manager module collects user personal information upon receiving an explicit consent from her.
2. The collected information is then transferred to the ad-selector module.
3. The ad-selector module creates a 'stereotype' of the user by putting her in a specific interest category based on her personal information. This category cannot be used to identify the user. PPTMA uses Google ad interest categories set [14].
4. The ad-selector module sends the user's stereotype to a specific ad-network.
5. The ad-network sends back a list of ads that might be of interest to users fall into the received stereotype. These ads will be stored in a local ad-pool.
6. Based on the user's personal information, the ad-selector module selects the most relevant ads from the ad-pool and show it to the user.
7. All user's views and clicks are recorded by the ad-selector module, and then reported to the ad-network for billing. The user's identity will not be exposed, instead, the ad-selector will use a pseudonym to refer to the user. This pseudonym will be changed after each billing report.

Figure 2 shows the framework of the PPTMA.

2.3 Federated Identities

Federated identities are simply a specific user's identities that can be implicitly combined so she can login into multiple websites using only one identity of them in a single sign-on process [15]. In the past few years, the most common adherence to this concept is the social login by which a user can use an existing identity from a social
networking service (e.g. Facebook, Twitter, etc.) to login into a third party website instead of creating a new login account (i.e. digital identity) specifically for that website [16].

The most profound specifications for a FIdM systems are the ones published under the name Liberty Alliance Project1. The name 'Liberty Alliance' has been changed to Kantara years after the specifications were published; however, the same specifications are still valid. Kantara Initiative2 focuses more on OAuth3, a claim-based identity management concept [15]. The whole scheme of Kantara is built on an open, standardised, communication framework (e.g. SAML SSO profiles, SOAP, etc.).

In a FIdM System there are three main parties:

1. **The Identity Provider or Identity Issuer (IdP)** this party issues an identity to the user, and is trusted by the other parties for the purposes of identity management.
2. **The Service Provider (SP)** the party that needs to identify the user before providing services to him/her.
3. **The User** who needs to use the SP services. Typically, the user employs a user agent (e.g. a web browser) as the means by which she/he interacts with the IdPs and SPs. Figure 3 shows the conceptual model of the FIdM system.

![Figure 3. The conceptual model of the FIdM system.](image)

The IdP and the SP have their own security policies. The IdP security policy includes information specific to individual users, including: how the user should be authenticated, which SPs it can send assertions about that user to, and which user attributes can be asserted. The SP security policy specifies which IdPs it trusts, how users must be authenticated by a specific IdP, and what types of attributes must be asserted by a specific IdP in order for a user to be granted the requested services.

A process called discovery of identity source (or simply discovery) [17] must take place during the user authentication process. This step enables the system to locate the IdP which is to be asked for an assertion. This step could be performed by either the user machine or the SP server; however, performing it on the user machine has the advantage of giving some protection against phishing attacks. Specifically, if a malicious SP performs discovery, then it could direct the user client to a fake IdP.

If there is a need for direct communication between the IdP and the SP, e.g. in order to exchange information about a user, then, depending on the identity management system in use, they may use a pseudonym or a temporary ID to refer to the user instead of the registered user identity. Such a procedure helps to preserve user anonymity.

If assertions are passed from the IdP to the SP via the user agent, then some identity management systems allow the user agent to prove its rightful possession of the assertion to the SP. This mitigates the risk of attacks in which an attacker uses an assertion issued to another user to impersonate that user. Such services are known variously as proof-of-rightful-possession, subject confirmation or proof-key methods. A variety of techniques for providing such proof have been proposed. Describing each technique would be out of the main scope of this paper.

In a Federated identity management system, the user might have one or more 'local' identities issued by SPs, in addition to a single identity issued by the IdP within a specific domain called a circle of trust (CoT). A typical CoT consists of a single IdP and multiple SPs. The IdPs of a CoT must be trusted by all the SPs within it. An SP can be a member of more than one CoT. A user can federate her/his IdP-issued identity with the local identities issued by SPs within the same CoT [17, 18].

As shown in Figure 4, the IdP and the SP may agree to use the same pseudonym to refer to a particular user, or they may use distinct pseudonyms. Regardless of how pseudonyms are used, it is clearly important that each party knows which pseudonym the other party will use to refer to a given user. For example, suppose that a user named Alice has three identities, an IdP-issued identity, Alice@IdP, and two local identities, Alice.1@SP1 and Alice.2@SP2, issued by, SP1 and SP2, respectively. The IdP could use one pseudonym (xxx, say) to refer to Alice when it communicates with SP1, and a different pseudonym (yyy, say) when it communicates with SP2. However, although the IdP is using the pseudonym yyy to refer to Alice when it communicates with SP2, SP2 may use a different pseudonym (y123, say) to refer to the same user when it communicates with the IdP. However, in some Federated identity management systems, the IdP and the SP do not agree on long-term pseudonyms for a particular user. Instead of using pseudonyms, the IdP and the SP use temporary IDs agreed during the authentication process. Such temporary IDs are typically only used for one working session.

---

1 http://www.projectliberty.org
2 https://kantarainitiative.org
3 http://oauth.net
The Liberty Alliance Identity Federation Framework (ID-FF) specifications [19] support identity federation and authentication (with SSO). They also describe the required techniques, including session management and identity/account linkage.

An ID-FF Liberty profile may best be defined as the combination of message content specifications and message transport mechanisms for a single type of client (that is, a user agent) [39]. There are many types of ID-FF Liberty profile, including SSO and Federation Profiles, Register Name Identifier Profiles, Identity Federation Termination Notification Profiles, Single Logout (or Single Sign-out) Profiles, Identity Provider Introduction, Namemlentifier Mapping Profile and NameIdentifier Encryption Profile. In this paper we are primarily concerned with the SSO and Federation Profiles; more specifically, the Liberty-enabled client and proxy (LEC) profile. Other the SSO and Federation Profiles include the Artifact profile, and the Browser POST profile [20].

2.3.1 ID-FF LEC Profile

This profile requires the involvement of a Liberty-Enabled User Agent (LEUA) in order to act upon the messages sent and received during the federation and authentication processes. An LEUA is typically implemented as a web browser enhanced with JavaScript components installed on the user machine, but it can be a software module installed on a mobile phone.

Figure 5 presents a sketch of the message flows within the ID-FF LEC profile in the case where the user has already been authenticated by the IdP. Note that the IdP can choose any authentication method according to its security policy. The Liberty-Enabling component must be installed on the user machine (or mobile) prior to the steps shown in the figure.

The message flows within the ID-FF LEC profile are as follows.
1. LEUA → SP : Log-in Request (HTTP Request with Liberty Enabled Header)
2. SP → LEUA : Authentication Request + 'optionally' an IdP List
3. LEUA or User : Obtains IdP
4. LEUA → IdP : Authentication Request
5. IdP → LEUA : Authentication Response + SAML Assertion
6. LEUA → SP : Authentication Response + SAML Assertion
7. SP → LEUA : Log-in Granted!

In step 1, the Liberty-enabling components add a special Liberty Enabled header to the HTTP request, so that the SP knows that the requesting user agent is Liberty-enabled. In step 2, the SP replies with a special Authentication Request message, which may include a list of trusted IdPs in addition to the SAML authentication assertion request. A brief example of this message is given in Figure 6.

The Liberty Specifications do not dictate how the user (or the LEUA) determines the identity of the IdP in step 3; this is left to the implementors of the Liberty-enabling
components. However, it is implemented, IdP discovery must be implemented on the user machine.

In step 4, the LEUA forwards the Authentication Request message to the IdP. Since the user has already been authenticated by the IdP, the IdP now sends a digitally signed Authentication Response message to the LEUA in step 5. This message is forwarded to the SP in step 6. Finally, the SP checks the forwarded Authentication Response message and, if it is acceptable, the user will be logged-in in step 7.

Messages in steps 4 and 5 must be carried over an SSL connection to provide confidentiality (integrity is guaranteed using an XML-Signature).

In Figure 7, we provide an example of a Liberty Authentication Response message. The message shown in the figure is tagged as <Lib:AuthResponse>, and it has a unique identifier (unique to the IdP). It also contains the unique identifier of the Liberty Authentication Request message (unique to the SP) to which it is a response. The issuer of the message is (http://IdP.com), i.e. the URL of the issuer IdP, and this functions as the identifier of the IdP. Similarly, the SP's URL is (http://SP.com), which functions as the identifier of the SP.

As stated above, this message contains a Liberty assertion (or <lib:Assertion>) which itself contains a Liberty authentication statement (or <lib:AuthenticationStatement>) which is an enhanced SAML authentication statement. The Liberty assertion has a unique identifier, and it also contains the unique identifier of the Liberty assertion request message to which it is a response. In addition, it specifies the duration of its validity; in the example it is valid between 11:32:49 on the first of January 2010, and 12:00:00 on the second of January 2010. Within the Liberty assertion is the <lib:Subject> element, which contains a shared pseudonym for the user. The assertion also contains the <SubjectConfirmation> element, which specifies the proof-of-rightful-possession method that has been used. Finally, the IdP's digital signature is included in the <ds:Signature> element.

3. Proposed Model

Conducting a PPTMA system via FIdM would result in a number of good improvements to both security and practicality of PPTMA. We suggest doing that by developing a minimal integration between PPTMA and Liberty ID-FF LEC profile. We have chosen this profile because it involves a smart user device that participates in the authentication (and advertising) process.

Before we describe the proposed integration model, we will briefly discuss the advantages of such integration to the PPTMA. Conducting PPTMA via FIdM will ensure:

1. **More realistic trust scheme**. In PPTMA the trust is scattered amongst the user, the user's mobile, the ad selector module, the developer, the ad network, the advertiser, and the third party app owner. Each party needs to trust all other parties which would not be a very realistic scenario. In our proposed model the trust will be focused on fewer parties; mainly: the IdP. The IdP is a trusted third-party by definition and it is responsible for creating (or at least registering) users' identities in a given CoT, along with approving and managing the authentication (and/or authorization) process in the federated identities echo system. The user, the user's device, and the SP all trust the IdP to manage one of the most (if not the most) critical security procedures. In PPTMA, this can be utilized to develop a simpler and more realistic trust scheme. Even the Ad-Selector itself can be placed in the IdP instead of the user's mobile.

2. **More efficiency and practicality**. This integration will minimize the computational and storage requirements on the user's mobile, since the Ad-Selector can be stored and operated on the IdP server. Pseudonyms, cryptographic mechanisms, secure messaging, strong authentication, and many more, are all already implemented in the FIdM system.

3. **More scalability and wider adoption**. Taking into consideration the growing number of users who use FIdM solutions for authentication (e.g.
social login), it is evident that PPMTA will gain a huge number of users.

4. **More trustworthy billing system.** The billing system will be encrypted by default with predefined pseudonyms, and can be stored and managed by the IdP.

In our proposed model the LEUA will be the user's mobile, the IdP will hold and manage the Ad-Selector and the billing modules, and the Ad-Network will be treated exactly as an SP.

We assume that the following integration steps take place after the user has authenticated herself to the SP (i.e. the Ad-Network) via the IdP using and FIdM system:

1. The user profile manager module, installed on the user's mobile, collects user personal information upon receiving an explicit consent from her.
2. The collected information is then transferred to the ad-selector module, installed on the IdP server.
3. The ad-selector module creates a 'stereotype' of the user by putting her in a specific interest category based on her personal information. This category cannot be used to identify the user.
4. The ad-selector module sends the user's stereotype to the SP (i.e. ad-network).
5. The SP sends back a list of targeted ads (received earlier from the advertisers) that might be of interest to users fall into the received stereotype. These ads will be stored in an ad-pool held by the IdP.
6. Based on the user's personal information, the ad-selector module selects the most relevant ads from the ad-pool and sends it to the user's mobile so it gets shown by the third party app.
7. All user's views and clicks are recorded on the mobile and reported to the ad-selector module at the IdP.
8. The ad-selector forwards the user's views and clicks report along with the bill to the SP (or ad-network). The user's identity will not be exposed, instead, the SP will use the same pseudonym used by the IdFM system during the authentication process to refer to the user. This pseudonym will be changed after each login. The billing report will be encrypted and digitally signed by the IdP using the same cryptographic mechanisms used during the authentication process.
9. The SP forwards the ad-clicks report and the bill to the advertisers for payment.

Figure 8 sketches the steps above.

### 4. Analysis

In this section, we provide a brief analysis of the proposed integration model.

#### 4.1 Security Threats

The security level of the integration model would be totally inherited from the security level of the FIdM system, which is evidently higher than the PPTMA. For example, FIdM uses SAML/SOAP protocols to convey exchanged messages, this means protection from attacks such as 'man-in-the-middle' or 'replay' because SAML messages contain nonces and time-stamps. In FIdM, all messages must be encrypted, using SSL/TLS for example, this ensures confidentiality. Same things apply when we consider techniques like 'levels-of-assurance' or 'proof-of-rightful-possession', or 'digital signatures'. All these effective security techniques are automatically inherited in the proposed integration model.

#### 4.2 Privacy

PPTMA offers a very good solution for the dilemma of preserving user privacy within a TMA system. The same solution is adopted in the proposed integration model. Storing and processing the user's personal information on the IdP server would boost up the efficiency of the system, and will not affect the privacy since the IdP is trusted by definition to perform and conduct more critical security procedures such as authentication, authorization, and audit. The use of temporary pseudonymous assures a good level of user privacy.

#### 4.3 Billing

The proposed integration model states that the generation of the billing report will be held at the IdP server, instead of the user's mobile. This will give it more credibility to this
5. Implementation

A proof-of-concept implementation was successfully conducted. Since what we propose in this paper is a 'high-level' integration, a small simulator of the ID-FF LEC profile was programmed and tested with real ads we got from multiple ad-networks. The main goal of the implementation was to prove that the built-in features in FIdM systems (e.g. cryptographical mechanisms and pseudonyms) can be readily utilized by any PPTMA; and we did. We used both SAML/HTTP and SAML/SOAP techniques, and both were successful. A more sophisticated integration software will be built in the future.

6. Future Work

We intend to design a more detailed integration, taking into consideration a number of FIdM systems. Moreover, we will examine the possibility of integrating other identity management models (e.g. Claim-based IdM) with PPTMA, and study whether or not that would be computationally and commercially feasible. Finally, there might be no need to use the 'ad-selector' when integrating Claim-based with FFTMA, since the 'identity-selector' of the Claim-based IdM could be able to carry-out the ad-selector's tasks; this is worth pursuing in a future research.

7. Conclusion

In this paper we have proposed a high-level integration model between FIdM and FFTMA. This integration should result in a number of advantages to the TMA ecosystem as discussed above. This integration should enhance the PPTMA security and boost-up its acceptance and scalability. A brief analysis of the proposed model along with our future work were also presented.

References


Waleed A. Alrodhan received his B.Sc. degree in Computer Sciences from King Saud University (2002), his M.Sc. degree (with distinction) in Information Security from Royal Holloway, University of London (2005), and his Ph.D. degree in Information Security from Royal Holloway, University of London (2011). Currently, he is the Dean of the College of Computer and Information Sciences at Imam Muhammed Ibn Saud University. His research interests include privacy, identity management, federated identity, single sign-on, and secure web-based protocols.