Chapter 7
Tracking and Fingerprinting in E-Business: New Storageless Technologies and Countermeasures

Károly Boda
Budapest University of Technology and Economics, Hungary

Ádám Máté Földes
Budapest University of Technology and Economics, Hungary

Gábor György Gulyás
Budapest University of Technology and Economics, Hungary

Sándor Imre
Budapest University of Technology and Economics, Hungary

ABSTRACT
Online user tracking is a widely used marketing tool in e-business, even though it is often neglected in the related literature. In this chapter, the authors provide an exhaustive survey of tracking-related identification techniques, which are often applied against the will and preferences of the users of the Web, and therefore violate their privacy one way or another. After discussing the motivations behind the information-collecting activities targeting Web users (i.e., profiling), and the nature of the information that can be collected by various means, the authors enumerate the most important techniques of the three main groups of tracking, namely storage-based tracking, history stealing, and fingerprinting. The focus of the chapter is on the last, as this is the field where both the techniques intended to protect users and the current legislation are lagging behind the state-of-the-art technology; nevertheless, the authors also discuss conceivable defenses, and provide a taxonomy of tracking techniques, which, to the authors’ knowledge, is the first of its kind in the literature. At the end of the chapter, the authors attempt to draw the attention of the research community of this field to new tracking methods.

DOI: 10.4018/978-1-4666-4181-5.ch007
INTRODUCTION

Many Website operators involved in end-user oriented e-business have an interest in monetizing their user base (e.g., by realizing as many clicks on their advertisements as possible). One way of achieving this goal is collecting diverse information about the user (or profiling her), the vehicles of which are technologies that can be used for identifying returning visitors, and those that infer sensitive information (such as browsing history) that users are not necessarily willing to disclose (e.g., purchase preferences). This controversial method may be a necessity for better serving the needs of online customers, but it often goes beyond user demands, and is used for business purposes against the will of the clients.

It is not surprising that certain users are concerned about pervasive profiling. In fact, this problem has been discussed from so many viewpoints in academia that not only effective technological, but also more or less effective legislative countermeasures have been implemented in order to mitigate the privacy risks arising from the proliferation of a subset of profiling-related technologies. That said, this chapter is inspired by the rest (primarily by fingerprinting attacks), for which neither defensive technology nor the law seems to be able to keep up with the pace of development dictated by pro-profiling actors, leaving users of the Web powerless against and vulnerable to them.

PROFILING AND USER PRIVACY ON THE WEB

There are several motives of profiling users on the Web. For instance, an enormous number of Web services can be accessed for free; however, contrary to how it looks (or is communicated), free access often comes with a greater sacrifice of user privacy, as many of these companies gain revenue from profiling-related activities, such as pursuing behavioral advertising or monetizing user profiles by other means. According to the report of IAB Internet Advertising Revenue Report (IAB, 2012), the advertising revenues set a new record at $8.4 billion in Q1 2012, clearly showing the significance and the growth of the advertising industry.

In two recent studies, Goldfarb & Tucker (January and May, 2011) concluded that targeted advertising can successfully influence individuals in favor of buying a product, and targeted advertising has a significant positive (economic) impact on advertising. These claims are consistent with the ever-increasing presence of Web tracking techniques (the most used tools for profiling) and the revenues the industry reached in 2012, presumably with a continuously increasing proportion of behavioral advertising. By summarizing the measurements of Krishnamurthy & Wills between 2006 and 2012 (Krishnamurthy & Wills, 2006; 2009; Krishnamurthy, 2010), Mayer & Mitchell (2012) highlighted that the coverage on top sites of large tracking companies increased during these years, and so did the number of trackers per page. Today, researchers estimate that there are trackers capable of monitoring more than one fifth of user activity while browsing online (Roesner et al., 2012).

However, as can be expected, tracking for targeted advertising is not favored by users. A Harris Interactive poll (Krane, 2008), a TRUSTe survey (2009), and a nationally representative telephone survey in the USA conducted by Turow et al. (2009) uniformly confirmed that the majority of individuals (around 60% in all cases) found it uncomfortable when they faced advertisements on Websites adjusted to their preferences by previously observing their online activities. A more recent online survey conducted by McDonald & Cranor (2010) also confirmed this result with 55% of respondents rejecting targeted advertising; another recent survey conducted by TRUSTe in partnership with Harris Interactive (2011) reported a higher rate of respondents, namely 85%, who