Electronic Commerce Research and Applications 15 (2016) 52–64

Contents lists available at ScienceDirect

Electronic Commerce Research and Applications

journal homepage: www.elsevier.com/locate/ecra

Shopping for privacy: Purchase details leaked to PayPal ☆

Sören Preibusch [a,*], Thomas Peetz [b], Gunes Acar [b], Bettina Berendt [b]

[a] Microsoft Research, UK
[b] KU Leuven, Belgium

ARTICLE INFO

Article history:
Received 25 January 2015
Received in revised form 1 November 2015
Accepted 23 November 2015
Available online 7 December 2015

Keywords:
Privacy
Online payments
Payment providers
PayPal
Tracking
Electronic commerce
Electronic retailing
Data minimisation
Data leakage

ABSTRACT

We present a new form of online tracking: explicit, yet unnecessary leakage of personal information and detailed shopping habits from online merchants to payment providers. In contrast to the widely debated tracking of Web browsing, online shops make it impossible for their customers to avoid this dissemination of their data. We record and analyse leakage patterns for the 881 most popular US Web shops sampled from actual Web users' online purchase sessions. More than half of the sites we analysed shared product names and details with PayPal, allowing the payment provider to build up fine-grained and comprehensive consumption profiles about its clients across the sites they buy from, subscribe to, or donate to. In addition, PayPal forwards customers' shopping details to Omniture, a third-party data aggregator with even larger tracking reach than PayPal itself. Leakage to PayPal is commonplace across product categories and includes details of medication or sex toys. We provide recommendations for merchants.

1. Introduction

1.1. Online payment providers process rich transaction data

Online payment handling is a key enabler for electronic retailing and a growing business opportunity as mobile commerce takes off. Contactless payments have been pioneered in successful yet isolated applications, such as public transport (e.g., Oyster in London, touch & travel in Germany) or entertainment (e.g., Disneyland Finextra Research 2009, Starbucks (Hamblen 2012). General-purpose digital wallets and near-field payment capabilities are now integrated in all major mobile phone operating systems (Google 2014, Microsoft 2014, Apple Inc. 2014) and promise wider adoption across verticals.

Payment providers are intermediaries between merchants and their customers who buy and then pay for goods and services. As intermediaries, payment providers necessarily gain insight into the transaction as they process personal information, just like the delivery company will need the customer's postal address. The minimum data requirements for payment handling are the order total, the receiving merchant and an authenticated payment instrument. This corresponds to data items traditionally collected during credit card transactions. However, a much richer set of data items becomes available for online, mobile and in-app purchases, including an itemised statement of the goods purchased or information about the buyer, allowing value-added services. Amongst credit card issuers, these data are known as Level II and III but have been rarely available for point-of-sale or transactions (Software Inc. 2014).

The move towards richer transaction details is driven and enabled by three factors: first, the extended role of payment providers as shopping cart solutions, so that itemised data availability becomes a necessity; second, technically enabled by the lack of data length restrictions found in legacy payment processing; third, the mining of detailed transaction data for fraud detection and prevention (Klarna 2013). For instance, MasterCard reported acceptance by over 19 million merchants worldwide back in 2001, but only 1% would be able to "capture and transmit Level II and Level III data". These include itemised product descriptions, quantities and prices (MasterCard 2001), but still fewer details than what new online payment providers collect.

☆ Preliminary results were presented as a short paper at the Financial Cryptography and Data Security 2015 conference. This article contains the full analysis.

* Corresponding author.
E-mail address: mail@soeren-preibusch.de (S. Preibusch).
URLs: http://preibusch.de (S. Preibusch), http://www.kuleuven.be (T. Peetz), http://www.kuleuven.be (G. Acar), http://www.kuleuven.be (B. Berendt).

http://dx.doi.org/10.1016/j.elerap.2015.11.004
1567-4223/

1.2. Potential benefits of data collection by payment providers

Fraud detection and prevention is the most-publicised benefit of collecting and inspecting purchase details. The rise of riskier card-not-present transactions over the Web or on mobile has mandated new efforts in fighting crime. Between 2002 and 2012, the most recent year for which data are available, the annual fraud losses on UK-issued payment cards has decreased from £427 million to £388 million. Whereas counterfeit, lost or stolen card fraud has decreased from £257m to £97m (−62%) during that period, card-not-present fraud for electronic commerce alone has quintupled from £28m to £140m and now accounts for the majority of losses (Financial Fraud Action UK 2013). Despite continued e-commerce growth, fraud volumes have been decreasing since their peak in 2008. The industry attributes these accomplishments to automated cardholder address verification and card security codes, to initiatives like MasterCard's SecureCode and Verified by Visa, and to the "effectiveness and sophistication of customer-profiling neural networks that can identify unusual spending patterns" (Financial Fraud Action UK 2013). The required collection of details about buyers and their purchases is therefore attractive for payment providers and merchants who can benefit from lower fees. As another example, the payment provider Klarna allows customers to pay after order placement and shipping. At the same time, it absorbs the credit risk for merchants and controls losses through risk assessment based on diverse factors, including purchase details (Gustafsson and Magnusson 2014).

Fighting payment fraud is only one of many more applications for purchase information. Payment providers have a twofold incentive to collect details for the transactions they process. One the one hand, they can use the additional data for operational efficiency in a broad sense; on the other hand, they can offer convenience features to consumers.

1.2.1. Operational efficiency

Payment providers operate in a highly regulated environment and some obligations cannot be fulfilled efficiently unless purchase details are known. They must comply with tax and legal requirements, such as products prohibited in certain regions (e.g., gambling, alcohol sales) or money laundering. They must also detect and prevent crime, such as fraud and policy violations. As an example of transaction monitoring, PayPal has "hundreds of highly trained specialists working around the clock to prevent fraudulent activity and identify suspicious transactions" (PayPal 2015). Details from past transactions are also a shared secret between the provider and its customers, and can be used for additional authentication or account recovery. Purchase details can be monetised for product innovation, as market research, and through direct marketing on an individualised basis. Insofar as payment providers provide escrow services and help buyers who have been defrauded by the merchant, transaction details can be used for risk screening. For instance, PayPal's buyer protection only covers certain physical goods. Whilst mainly in the self-interest of the provider, operational efficiency enables payment services for consumers and merchants at acceptable fees in the long run.

1.2.2. Convenience features

Buyers can enjoy peace of mind when their purchase details are displayed back to them in the very moment when making the payment. They can also inspect the transaction history in their account and get a detailed statement of previous purchases. When payment providers collect purchase details, they can offer sought-after spending reports and financial self-analysis.

1.3. Privacy concerns

The large-scale collection and processing of personal details causes privacy concerns. Concern is no longer limited to traditional items of personal information like address or demographics, but increasingly about consumption behaviour. Despite the quantified-self movement and although Web users volunteer personal information with high prevalence (e.g., 55% knowingly entered their weekly spending behaviour into a Web form where this item was optional, Malheiros et al. 2013), extended records of usage data are problematic. Widespread tracking of browsing patterns by Websites and aggregators has raised attention in mainstream media (WSJ Online 2013). Browsing history leaked to advertisers (TRUSTe 2009), electricity consumption recorded by smart meters (McDaniel and McLaughlin 2009), or mobility trajectories in pay-as-you-drive insurance policies (Scism 2013) have all been found to be associated with elevated privacy concerns. Of particular interest is shopping data, whose value is demonstrated through myriads of loyalty card schemes. Purchase tracking now happens across merchants and channels (online/offline) and even if users are not enrolled in a loyalty scheme (Valentino-DeVries and Singer-Vine 2012, Duhigg 2012).

Our research looks at the tracking capabilities of payment providers, namely PayPal. An illustrative example is provided in Figs. 1 and 4.

Our research motivation is the ability of payment providers to collect purchase details at scale. As in the domains of Web tracking and analytics, a small number of providers cover multiple Websites (merchants) and can link transactions across those. Compared to cookie-like tracking, the privacy issues are exacerbated:

- Embedded tracking code is—in principle—ancillary to the core functionality of the Web page and can safely be filtered out (e.g., with ad-blockers or Tracking Protection in Internet Explorer). Payment handling is however essential to shopping, and users cannot complete the transaction without interacting with the payment provider.
- Unlike browsing patterns linked to a cookie identifier, consumption patterns linked to a payment method are not pseudonymous but identifiable through offline details such as credit card numbers or bank account details, which often include full name.
- Payment cards or account information serve as persistent identifiers, allowing longitudinal linkage of multiple transactions even across different logins or accounts with the payment provider.
- Consumers are typically unable to evade such data collection unless they refrain from shopping with the given merchant. The collection of shoppers' details is a negative externality of the contract between the merchant and the payment provider.
- Payment handling is universal across sellers and sectors. Consumer details are collected and merged across transactions even for sensitive products and merchants. This includes pharmacies or adult entertainment, for instance, where shoppers deliberately moved out of the high street and onto the Web in a pursuit of privacy.

Privacy threats arise from detailed purchase patterns when more than the minimum data required are collected. The principle of data minimisation has long been codified in national law and international privacy guidelines, such as the "collection limitation principle" in the OECD privacy framework (OECD 2013) or the Madrid Privacy Declaration (The Public Voice 2009). The principle of data minimisation as such is now contained in the text of the European Union's upcoming General Data Protection Regulation (European Commission 2012; Council of the European Union