Impeding Individual User Profiling in Shopper Loyalty Programs

Philip Marquardt, David Dagon, and Patrick Traynor
Converging Infrastructure Security (CISEC) Laboratory
Georgia Tech Information Security Center (GTISC)
School of Computer Science, Georgia Institute of Technology
Atlanta, GA 30332
{pmarq, dagon, traynor}@cc.gatech.edu

Abstract. Shopper loyalty club programs are advertised as a means of reducing prices for consumers. When making a purchase, a customer simply scans their keyring tag along with the items they intend to buy and is granted a reduction in the total price. While the use of these cards results in a visible reduction in price, customers are largely unaware of the privacy implications of such discounts. In particular, the ability to link all purchases made by an individual customer allows retailers to develop detailed profiles that may reveal sensitive information, especially if leaked or sold to third parties. In this paper, we present ShopAnon, a mobile phone-based infrastructure designed to help consumers partake in shopper loyalty programs without allowing their transactions to be linked by a retailer. ShopAnon displays legitimate but random barcodes for specific retailers on each execution, and provides a number of operational modes that respond to the changing availability of resources and the specific privacy concerns of the user. Communications between the application and the database storing the barcodes occurs using an Oblivious Transfer protocol to prevent our system from exposing the barcode received by a requester. We design, implement and characterize the behavior of our application on the iPhone mobile platform, and demonstrate its practical efficiency (i.e., the ability to render random tags in less than 0.25 seconds via 802.11 links and approximately 3.9 seconds via a 3G cellular connection). Through this, we provide a powerful tool through which customers can improve their privacy in a retail environment.

1 Introduction

Businesses have launched a number of efforts in order to encourage customers to select themselves over the competition. Many stores offer a variety of indulgences (e.g., coffee stands, gourmet food, complimentary personal shoppers) in attempts to sell a "shopping experience" instead of just the individual items on their shelves. More recently, a large number of companies have tried to retain their customer base through the use of shopper loyalty programs. These marketing efforts are structured to provide regular shoppers with small discounts on their purchases once they become members of the program. Customers claim these discounts by presenting a scannable barcode, usually in the form of key tags/fobs, that links their purchases to their identity. Because presenting a membership tag results in a tangible savings, customers voluntarily sign

up for and participate in such programs without considering the potential consequences. In particular, by allowing a store to aggregate the purchases shoppers make, these programs enable businesses to profile and track their customers. As these profiles often contain sensitive information about individual customers (e.g., medications, products that may cause embarrassment) and may potentially be leaked or resold to third parties, these programs represent significant threats to consumer privacy.

In this paper, we present a means by which customers can participate in shopper loyalty programs with a greatly reduced risk of being profiled. We argue that a customer can be rewarded for their loyalty to a particular retailer without the wholesale surrender of their privacy. We achieve these ends through *ShopAnon*, an application designed for mobile phones that anonymously downloads loyalty club tags for its users. ShopAnon does not simply provide a trusted third party through which barcodes can be filtered. Instead, it allows users to automatically receive barcodes based on their current physical location while implementing an *Oblivious Transfer* protocol to prevent our infrastructure from learning the identity of the card served to the requester. Through these mechanisms, ShopAnon allows customers that wish to remain anonymous to be rewarded for patronizing the retailer of their choice. In so doing, we make the following contributions: 1) Develop an architecture capable of impeding the profiling of individuals in shopper loyalty programs, 2) Create a number of modes so as to provide consumers with operational robustness, 3) Implement, test and characterize our application on the iPhone platform.

Note that our system is robust against campaigns by retailers designed to identify and remove tags used in our system. Specifically, *we experimentally demonstrate that currently deployed systems rely on offline databases, allowing us to use cards for months after they have been deactivated.* Without the large-scale replacement of the cash registers and backend database systems with online systems, our approach will be able to assist users in protecting their privacy.

2 Threat Model

ShopAnon fundamentally relies upon the ability of consumers to use their mobile phone while making a purchase. There is no incentive for a store to support the ShopAnon program. Fortunately, other applications already allow users to store all of their shopper loyalty cards on their mobile phones [2]. These programs differ from ShopAnon in that the user remains traceable because the tag presented at each transaction is the same. However, the existence of these other programs allows customers to run ShopAnon without necessarily appearing to be avoiding profiling. If a retailer decides to prevent all such applications or devices from being used in their stores, our application is unlikely to be able to help protect the user. However, we believe that such a rule would be generally unenforceable and is more likely to irritate customers.

Retailers can download our software and potentially learn the values of some of the tags used in our system. Because of cost and highly distributed nature of current infrastructure, the databases responsible for processing shopper purchases function as an offline, batch-processing system. All updates to the system are completed by executing a bulk update during off-peak times. This results in the ability for the scanner to accept

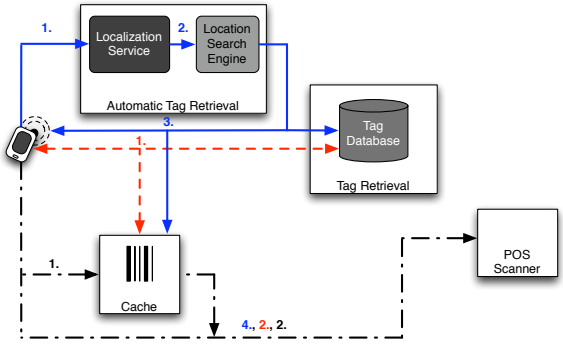

Fig. 1. A high-level overview of the ShopAnon architecture. The phone determines its location and queries a search engine for nearby businesses within a fixed radius. The response is filtered through the list of known shopper loyalty companies and then a request is sent to the tag database. The client then performs oblivious transfers with the database and receives the appropriate tag. Note that a client may skip the search phase by identifying the store manually (dashed red line) or by using a cached card (dot-dashed black line.) Barcodes are then displayed on the screen of the mobile phone and read by the Point-of-Sale (POS) scanner.

any shopper card that displays the correct barcode formatting for the store's shopper program. Supporting the instantaneous revocation of a card for any reason would require that these systems operate in an online fashion; however, such a change would be prohibitively expensive as it would require enabling all cash registers to access the database during each transaction involving a loyalty club card. Section 4.3 experimentally demonstrates the offline nature of current systems.

3 System Architecture

In this section, we present the details of the ShopAnon architecture. The *Automatic Protocol Execution* mode of operation attempts to provide a user with a suitable random barcode without any additional intervention. Automating such difficult tasks improves the usability of the application and increases the probability that the application will be accepted. The extended version of the paper provides more detail on modes of operation. Figure 1 provides a high-level overview of the ShopAnon architecture.

Location-Based Search Knowledge of one's physical location is a necessary precondition for automatically delivering relevant barcodes. Such information must be translated into the identity of a store before it can be used. We rely on location-based search engines (e.g., Google Maps, Yahoo! Maps, MapQuest) to assist in this process; however, there are a number of ways in which such systems can be used depending on the privacy concerns of the user [16, 11, 12]. In general, Shopanon will request a user's GPS coordinates and query the local-based search engine with a specific or general query provided by the user for finding stores within the specific search radius.

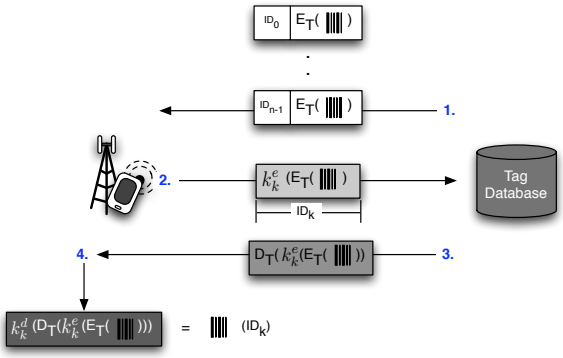

Fig. 2. A 1-out-of-n oblivious transfer as used in ShopAnon. The client 1) receives n encrypted barcodes from the Tag Database. Using the prepended identifier, the client determines which tag it would like to receive, 2) removes the identifier, encrypts the tag with its key (modular multiplication with random k) and returns it to the tag database. The Tag Database then 3) decrypts the received tag, and sends the resulting ciphertext to the client. Finally, 4) the client decrypts the received ciphertext (modular division by k) and recovers the desired barcode. Note that the Tag Database can not determine the identity of the tag selected by the client because it can not fully decrypt any of the ciphertexts.

3.1 Tag Retrieval

Preventing a customer from being tracked by a loyalty rewards program has little practical value if we simply shift responsibility for tag dissemination to a "trusted third party". For instance, nothing prevents such an entity from changing their terms of service and eventually selling the data collected regarding user request patterns. Moreover, a system built upon this assumption fails to prevent insiders, intruders or those with access to server logs from similarly exfiltrating such data. We aim to provide a service in which correlation between a user and a tag is difficult for *anyone* except the user.

We address these concerns through the use of *Oblivious Transfer* (OT). This cryptographic primitive allows a server Alice to offer a receiver Bob k-out-of-n pieces of data without allowing Alice to determine the identity of the k pieces delivered to Bob while preventing Bob from learning more than k pieces of data. *This approach is advantageous as it not only prevents our server from learning the actual tag delivered to a mobile for a given transaction, but also because it prevents an adversary capable of compromising the server from using log data to learn about a client's past behaviors.*

We selected an OT scheme proposed by Huang et al. [7]. This work is based on RSA and relies on the security guarantees of this algorithm. Unlike the above definition, Alice now sends Bob a series of messages $m_0, \cdots, m_{n-1}$. Alice encrypts these messages such that $X_0 = m_0^e, \cdots, X_{n-1} = m_{n-1}^e$. Bob then selects k of the n received messages, selects k random numbers and encrypts the selected subset to create the set $Y_0 = X_{k0} \cdot k_0^e, \cdots, Y_{k-1} = X_{kk-1} \cdot k_{k-1}^e$. Alice receives the k messages from Bob, which she can not uniquely identify from the set M, and decrypts them such that $C_0 = Y_0^d, \cdots, C_{k-1} = Y_{k-1}^d$. Bob then receives these messages, removes his secret and is able to retrieve the contents of the k selected messages. Figure 2 provides an overview of the 1-out-of-n scheme used in this work.

0. $C \rightarrow D : V_0, \cdots, V_{n-1}$
1. $D \rightarrow C : V_0 | \langle B_{V_{0,i}}, t \rangle_{k_D^+}, \cdots, V_{n-1} | \langle B_{V_{n-1,i'}}, t' \rangle_{k_D^+}$
2. $C \rightarrow D : \langle B_{V_{j,i''}} \rangle_{k_D^+} * Y$
3. $D \rightarrow C : \langle \langle B_{V_{j,i''}} \rangle_{k_D^+} * Y \rangle_{k_D^-}$
4. $C \rightarrow V_j : B_{j,i''}$

C, D, V : Client, Database Server, Vendor
B, Y : Barcode, Client secret value
k_D^+, k_D^- : Public and private key for Database D
$\langle B \rangle_{k_D^+}$: B encrypted with public key k for D

Fig. 3. The ShopAnon protocol and variable definition. We formally define the interaction between the application and the Tag Database. Each message corresponds to the messages in Figure 2. Note that Message 0, the initial request, is not included in the previous figure.

The scheme proposed by Huang et al. is attractive for a number of reasons. First, the intuitive nature of this scheme makes its implementation relatively straightforward. Moreover, unlike related OT schemes based on Diffie-Hellman, we can parameterize our implementation to be very efficient based on small values for e in RSA while remaining resistant to known small exponent attacks. For a k-out-of-n system based on this cryptosystem, a client is required to perform $4k$ modular multiplications, whereas the server performs k modular exponentiations and $2n$ modular multiplications. Given that other related schemes generally require $O(kn)$ modular exponentiations by the server and $O(k)$ modular exponentiations by the client [5, 9, 13, 15], the selected scheme is more appropriate for potentially constrained mobile devices.

Our protocol is specified formally in Figure 2.

4 System Evaluation and Analysis

4.1 Experimental Setup

We used the following development, software, and hardware environments to develop and test the ShopAnon application. The client application was developed for the iPhone 3G OS v3.1.2. Xcode v3.1.2 was used as the development environment for both the client and server applications. During the development phase, iPhone simulator v3.1 was used for verifying correct functionality. All development software was run on an Apple MacBook Pro with 2.2GHz Intel Core 2 Duo processor, 4 GB 667 MHz DDR2 SDRAM, running Mac OS X v10.6.2 (Snow Leopard). The ShopAnon application is written in Objective-C with some C functionality and compiled using GCC v4.2. We cross-compiled and used GMP v4.3.1 for cryptographic functions and pseudorandom number generation. Figure 4 shows the working application.

4.2 Microbenchmarks

Our first set of experiments ran the tag retrieval portion of ShopAnon over an 802.11g wireless connection. We ran a total of 10 experiments for each of the above data points and calculated 95% confidence intervals of at least one order of magnitude smaller than

Fig. 4. The ShopAnon user interface. In these images, the user selects "Find By Store Name", selects a particular retailer and automatically receives one of the barcodes stored for that vendor in the Tag Database.

the mean. ShopAnon rendered the received barcode in an average of 0.229 ± 0.0166 seconds after selecting the vendor for which a barcode was needed. The cryptographic overhead related to this delivery was also extremely low - the application was able to perform encryption and decryption in 0.023 ± 0.002 seconds and 0.008 ± 0.0002 seconds, respectively. Accordingly, our fears of significant overhead in order to perform OT operations were alleviated. The performance of our protocol over the 3G UMTS cellular data link was significantly different from 802.11g. All network operations conducted via the cellular link were at least an order of magnitude more expensive than WiFi, with a total protocol execution time of 3.939 ± 0.560 seconds. These results are expected based on the architecture of the network. In particular, the high cost of connection establishment [14] and the use of scheduling algorithms such as Proportional Fair that favor short high bandwidth bursts as opposed to uniform distribution of traffic with low latency [6].

Running ShopAnon on a mobile phone is practical given these experimental observations. However, we must still determine whether or not the barcode values delivered by these exchanges are usable by current scanning technology.

4.3 Field Testing Experience

Card Collection Experience The barcodes made available in ShopAnon represent real loyalty club card accounts. We have not attempted to reverse-engineer the algorithms used to generate customer identifiers and have therefore not included any "forged" cards in our system. Instead, individuals in our research group visited a range of stores offering these cards and retrieved them. Our experiences were mixed, even across multiple locations within the same chain. Whenever possible, group members requested receiving tags without any name associated with them. While some of the stores we visited provided such cards without question, the majority of stores required that we provide some identifying information. Most employees, however, knowingly allowed us to enter false information. This reaction was expected and has been observed by other parties concerned about privacy in this space [3]. Only one out of the five chains we visited required that we provide a state-issued identification card, such as a driver's license.

However, we were able to acquire a card from the same chain online without the need to provide this verifiable information. At no time were our requests for cards rejected. Moreover, regardless of the information that was entered at the time of activation, all of the cards that we have collected continue to allow us to receive the discounts at each of the respective stores. We use these collected cards to populate our Tag Database and, as mentioned above, plan to allow users to submit their own cards to the system as well.

Card Invalidation Experience To gain insight on how the barcodes are processed during the checkout time, we chose two barcodes from two different retailers and tried to manually invalidate them. Our hypothesis is the POS systems are not connected to an online central database, so either the cards can not be deactivated or all deactivated cards are downloaded by all stores periodically making deactivation processing times lengthy. Such invalidation problems are well known in the field of certificate revocation [8].

Retailers were contacted and provided with the barcodes requiring invalidation. One operator informed us the deactivation process would take four to six weeks while another operator indicated immediate processing. As expected, both cards were tested for six months and continued to succeed in receiving discounts to qualified purchases. This testing proves that the system being used to process discount cards is completely offline, validating the assumption made in our threat model. Given that the cost of retrofitting all of the cash registers in a shopping chain and the backend database to support an system capable of processing thousands of transactions per second is very high, ShopAnon is resistant to deactivation attacks. While vendors may eventually purchase such highly-capable infrastructure, our solution is an important first step in protecting consumer privacy in this space.

Application Usage Results Having collected a number of valid shopper loyalty club tags and entered them into the Tag Database, we then sought to determine how well such tags could be read by barcode scanners. While other tag storage applications have already shown that it is possible to read barcodes from an iPhone screen [2], users regularly complain that such applications suffer from poor detection and accuracy. These tests accordingly attempt to better quantify such issues. We note that like many other mobile phones, the iPhone tested in these experiments used a scratch-resistant plastic film on the screen. Our experiments were conducted both with and without this cover.

Our attempts to detect barcodes were extremely successful with handheld scanning guns. We specifically ran this set of tests using the Metrologic Ms1690 Focus handheld scanner, which is also used by a number of major retailers on a national scale. In these tests, barcodes were readable ten out of ten times when the screen protector was removed; however, our results were not as consistent when the screen protector was in place with only seven out of ten attempts being successful. We also tested the performance of our barcodes against a UniComp PCT2 Price Checker and found that these devices were able to read the barcode rendered by the ShopAnon application nine out of ten times when the screen protector was in place and ten out of ten times when the screen was unobscured.

We also measured traditional flatbed barcode scanners by testing the NCR RealPOS High Performance Bi-Optic Scanner/Scale [10], which is used by major chain stores

throughout the country. In spite of multiple attempts from a variety of angles, orientations, distances and in the presence and absence of the screen protector, we were unable to scan the barcode displayed by ShopAnon. This was expected given previously documented complaints [1, 4]. Fortunately, most of these systems come with an optional hand scanner, which a customer can request their cashier use on the displayed barcode.

5 Conclusion

Shopper loyalty club programs offer their members demonstrably lower prices than those given to non-members. Unfortunately, in the pursuit of these savings, users often unknowingly sacrifice their privacy and allow themselves to be profiled. In this paper, we develop ShopAnon, an infrastructure that allows consumers to receive the benefits associated with such programs while allowing their transactions to remain unlinkable. Using an application running on their mobile phone, consumers download random legitimate barcodes to be presented at the checkout. Such barcodes can be downloaded automatically or with manual assistance and cached to allow for potential offline operation in the future. To demonstrate that this system is practical, we design, implement, measure and field test our application on the iPhone platform and show that users can execute this protocol in a matter of seconds (approximately 3.9 seconds over 3G UMTS links and 0.2 seconds via 802.11g). We then provide a number of options by which more advanced users can further limit their exposure. In so doing, we provide a robust tool and an important first step by which consumer privacy can be protected in a retail environment.

References

1. AppShouter. CardStar iPhone App Review. `http://appshouter.com/iphone-app-review/iphone-app-review-cardstar/`, 2009.
2. CardStar, Inc. CardStar. `http://www.mycardstar.com/`, 2009.
3. Consumers Against Supermarket Privacy Invasion and Numbering (CASPIAN). Is Big Brother in Your Grocery Cart? `http://www.nocards.org/`, 2009.
4. B. Fischer. AppAdvice: CardStar. `http://appadvice.com/appnn/2009/02/review-cardstar/`, 2009.
5. B. Gilles, C. Claude, and W. Stefan. Oblivious Transfers and Privacy Amplification. *Journal of Cryptology*, 16(4), 2003.
6. H. Holma and A. Toskala, editors. *HSDPA/HSUPA for UMTS*. John Wiley & Sons, Ltd, 2006.
7. H.-F. Huang and C.-C. Chang. A new design for efficient t-out-n oblivious transfer scheme. *Advanced Information Networking and Applications, International Conference on*, 2:499–502, 2005.
8. P. McDaniel and A. Rubin. A Response to 'Can We Eliminate Certificate Revocation Lists?'. In *Proceedings of Financial Cryptography*, 2000.
9. M. Naor and B. Pinkas. Efficient Oblivious Transfer Protocols. In *Proceedings of SIAM Symposium on Discrete Algorithms (SODA)*, 2001.
10. National Cash Register (NCR). NCR RealPOS High Performance Bi-Optic Scanner/Scale. `http://www.ncr.com/products_and_services/point_of_sale/pos_scanners/index.jsp`, 2009.
11. R. A. Popa, H. Balakrishan, and A. J. Blumberg. Vpriv: Protecting privacy in location-based vehicular services. In *Proceedings of the USENIX Security Symposium*, 2009.
12. P. Shankar, V. Ganapathy, and L. Iftode. Privately Querying Location-based Services with SybilQuery. In *Proceedings of the International Conference on Ubiquitous Computing*, 2009.
13. J. Stern. A New and Efficient All-Or-Nothing Disclosure of Secrets Protocol. In *Proceedings of the nternational Conference on the Theory and Application of Cryptology and Information Security (ASIACRYPT)*, 1998.
14. P. Traynor, P. McDaniel, and T. La Porta. On Attack Causality in Internet-Connected Cellular Networks. In *Proceedings of the USENIX Security Symposium (SECURITY)*, 2007.
15. W. Tzeng. Efficient 1-out-n oblivious transfer schemes. In *Proceedings of the Workshop on Practice and Theory in Public-Key Cryptography (PKC)*, 2002.
16. G. Zhong, I. Goldberg, and U. Hengartner. Louis, Lester and Pierre: Three Protocols for Location Privacy. In *Proceedings of the 7th Workshop on Privacy Enhancing Technologies*, 2007.