A Fuzzy Commitment Approach to Privacy Preserving Behavioral Targeting

Shantanu Rane
PARC, Inc.
3333 Coyote Hill Road, Palo Alto, CA.
srane@parc.com

Ersin Uzun
PARC, Inc.
3333 Coyote Hill Road, Palo Alto, CA.
euzun@parc.com

ABSTRACT

This paper describes a privacy-preserving framework for delivering coupons to users that approximately satisfy a predefined behavioral profile. The framework is designed to be non-interactive, i.e., vendor-side communication occurs only when it pushes coupons out to the users that it regards as potential customers. User privacy is protected by performing all targeting operations on the end-user's device. The protocol is based on a fuzzy commitment primitive that is realized using error correcting codes. The central idea is that a user is able to extract the coupon if her behavioral profile approximately matches the vendor's target profile. Unless the coupon is redeemed, the vendor discovers no information about the user's behavioral profile. The error correction coding framework enforces a natural tradeoff between the privacy of the vendor and the specificity of targeting. In other words, if the vendor wants to target a broad class of potential customers, it must reveal more information about its targeting strategy to ineligible users. Conversely, if the vendor wants to reveal less information about its targeting strategy to ineligible users, then it must target a more focused class of potential customers.

Categories and Subject Descriptors

K.4 [**Computers and Society**]: Electronic Commerce—*Distributed commercial transactions, Payment schemes*; K.4 [**Computers and Society**]: Public Policy Issues—*Privacy*

General Terms

Algorithms, Security, Privacy

Keywords

Fuzzy commitment; private behavioral targeting

MOBICOMM SPME 2014 Hawaii, USA

http://dx.doi.org/10.1145/2646584.2646591.

1. INTRODUCTION

Online advertising is becoming increasingly dependent on behavioral targeting of consumers. Vendors seek to customize their messaging based on data that they gather about a user's online behavior, physical context such as location and time of day, as well as social context such as the presence of other people of the same age in the vicinity [1]. The intended goal of behavioral targeting is to provide personalized and relevant messages to potential customers, while minimizing the number of irrelevant or annoying advertisements. However, indiscriminate data gathering of the sort that is practiced today, can potentially undermine the privacy of users [2, 3]. Therefore, it is necessary to develop methods that enable personalized messaging without compromising personal privacy.

There is an added concern, from the point of view of vendors. If targeted advertising is performed in the clear, e.g., by directly comparing a user's behavioral profile with a vendor's target profile and sending coupons or deals to matching customers, there is an incentive for users to cheat. Ineligible users may take advantage of the information contained in the vendor's target profile and attempt to redeem coupons. Thus, it is in the vendor's interest to reveal the coupon only to eligible users, while leaking little or no information to ineligible users. It is the goal of this paper to propose a solution that addresses privacy concerns of end-users and vendors, while still realizing the benefits of targeted advertising.

The proposed solution is non-interactive in the sense that the vendor pushes the coupon and its targeting strategy in encrypted form to the user, and then engages in no further communication with the user. Only eligible users, i.e., users whose behavioral profile approximately matches the target profile, are able to discover the vendor's targeting strategy and to decrypt the coupon. The user's behavioral data and any computational manipulation required to access the coupon is restricted to her device, thereby protecting her privacy. When a user redeems a coupon, the vendor discovers that her behavioral profile was indeed close enough to the target profile. If a user is not eligible, or is eligible

but chooses not to redeem a coupon, the vendor discovers no information about her behavioral profile.

The proposed solution is inspired by the method of protecting biometric templates by means of fuzzy commitment [4,5]. This is a method of enabling access control without storing a biometric in the clear at the access control device. In fuzzy commitment, a codeword is generated from a secret key and then perturbed by the user's biometric at enrollment. In this way the key is bound to the biometric, and stored in such a way that neither the key nor the biometric is easily accessible to an adversary. A user with a legitimate test biometric that approximately matches the enrollment biometric, can then extract the key and gain access. In our basic protocol, a vendor commits a secret key that can only be extracted by an eligible user, and later used to decrypt and redeem a coupon. We also present an enhanced protocol in which user authentication is enforced by a two-factor variant of fuzzy commitment [6], where the key-binding operation also incorporates a unique user ID. To our knowledge, this is the first work in which a vendor exploits fuzzy commitment to target users whose online behavior matches a target profile.

The remainder of this paper is organized as follows: Section 2 presents the problem setting, clarifies notation and describes the proposed non-interactive protocol that implements behavioral targeting. We present an embodiment based on error correcting codes, though other methods are possible. Privacy guarantees provided by the protocol are discussed in Section 3, along with an information-theoretic assessment of vendor privacy. In Section 4, we compare and contrast the proposed protocol with related work in behavioral targeting, before providing some concluding remarks.

2. BEHAVIORAL TARGETING FRAMEWORK

2.1 Problem Setting and Notation

We assume that the vendor's targeting strategy and the user's behavioral profile can be represented as vectors $\mathbf{v}$ and $\mathbf{u}$ in the same n-dimensional space. Such a representation can readily be obtained by recording pieces of information about a user's activity that she has consented to share while browsing the vendor's website. This approach has been suggested for generating user-specific responses to search queries [7]. For our application, elements of the behavioral vector would be values of attributes such as the number of relevant search queries during the past week, number of similar products purchased in the past month, number of friends who have purchased the items being queried, percentage of time spent looking at a certain class of products, and so on. We assume that the schema, i.e., the ordering of attributes in $\mathbf{u}$ and $\mathbf{v}$ is the same, so that the distance between the two vectors is a meaningful indicator of their similarity.

For the purposes of exposition, assume further that the elements of these vectors are binary, representing, for instance, yes/no answers to a variety of questions about consumer behavior. Thus, $\mathbf{u}, \mathbf{v} \in \{0,1\}^n$. Our construction is based on algebraic error correcting codes, so the generalization to larger finite fields can be performed in a straightforward manner, allowing the behavioral profile to contain non-binary attributes such as the examples cited above. The user is uniquely identified in the vendor's database by an n-dimensional binary vector $\mathbf{z}$ that serves as her user ID.

The vendor and user have agreed that they will use an error correcting code, that is defined by a generator matrix $\mathbf{G} \in \{0,1\}^{m \times n}$. For the purpose of encrypting the coupon, the vendor uses a binary key $\mathbf{k} \in \{0,1\}^m$, where $m < n$. To determine whether the user is eligible to redeem the coupon or not, she needs a way to verify whether she has correctly extracted $\mathbf{k}$. For this purpose, the vendor and user employ a cryptographic hash function, such as SHA2, which we denote by $h(\cdot)$.

2.2 Basic Fuzzy Commitment-Based Protocol

The protocol below describes a first step in which the vendor generates an encrypted coupon, and a second step in which a user verifies whether she is eligible to redeem it. The two-party protocol proceeds as follows:

1. The vendor sends the following items to the user: $\mathrm{Enc}_k(\texttt{Coupon}\|\texttt{Nonce}\|\texttt{Signature})$, $h(\mathbf{k})$, and the n-dimensional target vector $\mathbf{t} = \mathbf{v} \oplus \mathbf{kG}$. Here, `Coupon` denotes the content of the delivered coupon, `Nonce` denotes a random number, and $\texttt{Signature} = \mathrm{Sig}_{\mathrm{vendor}}(\texttt{Coupon}\|\texttt{Nonce})$ is evidence that the vendor has digitally signed the encrypted transmission. The operation $\oplus$ represents modulo-2 addition. Vendor-side operations are depicted in Fig. 1.

2. The user first computes $\mathbf{t} \oplus \mathbf{u} = (\mathbf{v} \oplus \mathbf{u}) \oplus \mathbf{kG}$. The result is an errorprone version of the codeword $\mathbf{kG}$, where the errors are caused by the difference between the target profile $\mathbf{v}$ and the user's profile $\mathbf{u}$. As shown in Fig. 2, she performs error correction decoding to recover a candidate message $\mathbf{k}'$. If $h(\mathbf{k}') = h(\mathbf{k})$, she uses $\mathbf{k}$' to decrypt the coupon. If $h(\mathbf{k}') \neq h(\mathbf{k})$, she is unable to decrypt the coupon.

3. If the user has successfully decrypted the coupon and decides to redeem it, she can check its validity by verifying the vendor's signature.

2.3 Protocol with User Authentication

The protocol below describes an enhancement that prevents unauthorized users from decrypting the coupon. This is achieved by including the user ID $\mathbf{z}$ in the key-binding step as follows:

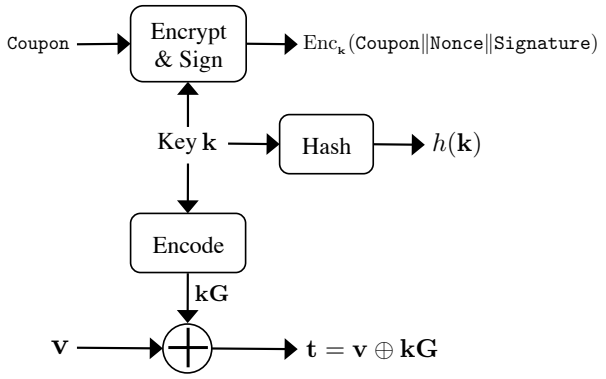

Figure 1: Vendor encrypts the coupon with a key k. It also binds the key to its target profile v using fuzzy commitment.

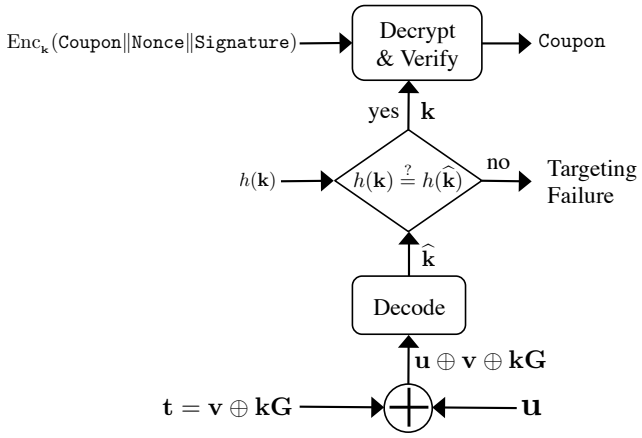

Figure 2: User attempts to recover the bound key k. If successful, she can decrypt the coupon.

1. The vendor sends the following items to the user: $\text{Enc}_k(\texttt{Coupon}\|\texttt{Nonce}\|\texttt{Signature})$, $h(\mathbf{k})$, and the n-dimensional target vector $\mathbf{t} = \mathbf{v} \oplus \phi(\mathbf{z}) \oplus \mathbf{kG}$. Here, $\phi(\cdot)$ is a pseudorandom function taking values in $\{0,1\}^n$. All other quantities are defined as before.

2. The user computes $\mathbf{t} \oplus \phi(\mathbf{z}) \oplus \mathbf{u} = (\mathbf{v} \oplus \mathbf{u}) \oplus \mathbf{kG}$. She then proceeds as in the basic protocol.

3. PROTOCOL ANALYSIS

3.1 User Privacy

In the above protocol, after the vendor has pushed the encrypted coupon to the user, there is no further communication between the vendor and the user. So long as the user does not redeem the coupon, the vendor finds out nothing about her behavioral profile. Only if the user successfully decrypts the coupon and redeems it, does the vendor discover that the targeting was successful, i.e., that the user's behavioral profile $\mathbf{u}$ is sufficiently similar to the target profile $\mathbf{v}$. The degree of similarity is determined by the parameters of the error correcting code. For example, if $\mathbf{G}$ generates a (n, k, d) block error correcting code, then a successfully redeemed coupon reveals to the vendor that, $\mathbf{u}$ differs from $\mathbf{v}$ in at most $\lfloor \frac{d-1}{2} \rfloor$ bits with high probability.

3.2 Vendor Privacy

The privacy of the vendor depends on whether the targeted user is eligible or ineligible. We consider these two scenarios from the basic protocol, as follows:

1. If the user is ineligible, the error sequence $\mathbf{v} \oplus \mathbf{u}$ cannot be corrected by the code $\mathbf{G}$. Hence, she cannot recover $\mathbf{k}$ and cannot decrypt the coupon. However, she still discovers some non-trivial information about $\mathbf{v}$ because the space of codewords $\mathbf{kG}$ has only m degrees of freedom, whereas the vector $\mathbf{t} = \mathbf{v} \oplus \mathbf{z} \oplus \mathbf{kG}$ has $n > m$ bits. The number of bits leaked about $\mathbf{v}$ depends on the design of $\mathbf{G}$, which in turn, depends on the targeting strategy. We elaborate further in the next sub-section.

2. If the user is eligible, the error sequence $\mathbf{v} \oplus \mathbf{u}$ is successfully corrected by the code $\mathbf{G}$, so she can extract $\mathbf{k}$ and decrypt the coupon. An eligible user discovers that the targeting has been successful. Furthermore, since she has correctly recovered $\mathbf{k}$, she can compute the vendor's target profile exactly using $\mathbf{t} \oplus \mathbf{kG} = \mathbf{v}$. If this is not desirable, then the vendor can anonymize its strategy by adding noise to $\mathbf{v}$. The statistical distribution of the noise would be chosen such that it provides some privacy against eligible users but still allows decoding of $\mathbf{k}$ via the code $\mathbf{G}$. We will return to this point in the analysis presented in the next sub-section.

Next, we briefly consider the effect of incorporating user authentication. Ideally, if the user ID $\mathbf{z}'$ differs from the targeted ID $\mathbf{z}$ even in a single bit, the user should not discover anything about the vendor's target profile $\mathbf{v}$, thus remaining oblivious of the targeting strategy selected by the vendor for user $\mathbf{z}$. By employing $\phi(\mathbf{z})$ (instead of $\mathbf{z}$) in the protocol with user authentication, we ensure that given $\mathbf{t} = \mathbf{v} \oplus \phi(\mathbf{z}) \oplus \mathbf{kG}$, an unauthorized user with ID $\mathbf{z}'$ cannot extract information about $\mathbf{v}$ unless $\phi(\mathbf{z}') = \phi(\mathbf{z})$. Since $\phi(\cdot)$ is a pseudorandom function, the equality will occur with exceedingly low probability. Note that this privacy guarantee applies to *semi-honest* users who follow the protocol as described. Vendor privacy would be compromised, for example, if an eligible user shared the key $\mathbf{k}$ and her user ID $\mathbf{z}$ with an ineligible user.

3.3 Information-Theoretic Analysis

We now elaborate on the two claims about vendor privacy in the case of eligible and ineligible users in the

basic protocol. This aspect is of interest not only because it differs qualitatively from vendor privacy guarantees encountered in previous work on behavioral targeting, but because fuzzy commitment appears to enforce a tradeoff between the specificity of the targeting strategy and the information revealed to ineligible users.

For information-theoretic analysis, let $P_{\mathbf{K}}$, $P_{\mathbf{U}}$, and $P_{\mathbf{V}}$ be the probability distributions from which are drawn the key, the user's behavioral profile and the vendor's target profile. Denote random variables having these distributions using uppercase letters $\mathbf{K}$, $\mathbf{U}$ and $\mathbf{V}$ respectively. Thus, the quantities $\mathbf{k}$, $\mathbf{u}$ and $\mathbf{v}$ in Figs. 1 and 2 can be viewed as specific values taken by these random variables. As indicated in our problem setting, these are all binary vectors. We make a reasonable assumption that the key $\mathbf{K}$ consists of i.i.d. Bernoulli-0.5 bits chosen independently of $\mathbf{V}$ and $\mathbf{Z}$. We make another simplifying assumption, which may not hold in practice, that the bits of $\mathbf{V}$ (and similarly $\mathbf{U}$) are i.i.d. Bernoulli-0.5. With this setup, we are first interested in the number of bits revealed to an *ineligible* user, when she receives $\mathbf{T} = \mathbf{V} \oplus \mathbf{KG}$ as described in Section 2. As in the case of biometric vectors [6], the privacy leakage is evaluated in terms of the mutual information as:

$$\begin{aligned} I(\mathbf{V};\mathbf{T}) &= H(\mathbf{T}) - H(\mathbf{T}|\mathbf{V}) \\ &= H(\mathbf{KG} \oplus \mathbf{V}) - H(\mathbf{KG}) = n - m \end{aligned}$$

For a block error correcting code $\mathbf{G}$, $n - m$ is indeed the number of error correction bits needed to extract $\mathbf{K}$ given the noisy codeword $\mathbf{V} \oplus \mathbf{U} \oplus \mathbf{KG}$, where the noise is given by $\mathbf{V} \oplus \mathbf{U}$. This points to an interesting tradeoff between the privacy of the vendor and specificity of its targeting strategy, which we explain below.

Concretely, if the vendor wants a broad class of its users to access the coupon, then it must allow different kinds of users to correct the noisy bits given by $\mathbf{V} \oplus \mathbf{U}$. This means that, it should employ a code $\mathbf{G}$ with strong error correction capability, i.e., with a large number of parity bits, $n-m$. Thus, in this scenario of loose targeting, the vendor must reveal more information about the target profile to ineligible users. On the other hand, if the vendor wants to target a specific class of users, then it needs to allow only that specific group to correct the noisy bits given by $\mathbf{V} \oplus \mathbf{U}$. This means that, it should employ a code $\mathbf{G}$ with weak error correction capability, i.e., with few parity bits, $n - m$. Thus, in this scenario of specific targeting, the vendor reveals less information about the target profile to ineligible users.

Finally, recall from the discussion about vendor privacy above that, in order to anonymize the vendor's strategy against *eligible* users, the vendor can add noise to $\mathbf{v}$. The stronger the noise used for anonymization, the larger the number of parity bits, $n - m$, needed for error correction. Thus, using a similar argument as above, the vendor faces a tradeoff between achieving privacy against eligible users versus achieving privacy against ineligible users.

3.4 A Note on Robustness

A user should be able to decrypt the coupon with high probability if her behavioral vector $\mathbf{u}$ is sufficiently close in Hamming distance to the vendor's target profile $\mathbf{v}$. In general, for a (n, m, d) error correcting code, the user is able to correctly extract $\mathbf{k}$ when the number of symbol errors w between $\mathbf{u}$ and $\mathbf{v}$ is less than $t := \lfloor \frac{d-1}{2} \rfloor$, which is the error correction capability of the code. When the number of errors is $w > t$, the system will either commit a decoding error, i.e., decode a wrong message $\mathbf{k}' \neq \mathbf{k}$, or report a decoding failure, i.e., not find any codeword within a distance t of $\mathbf{u} \oplus \mathbf{v} \oplus \mathbf{kG}$. The case of decoding error is disambiguated with high probability using the cryptographic hash comparison shown in Fig. 2. The case of decoding failure can be addressed using list decoding [8], which effectively captures codewords beyond a distance of t, and subjects them to the hash comparison. The size of the list should be chosen such that the system will reliably allow a high percentage of eligible users to decrypt coupons, while restricting the percentage of "lucky" ineligible users to a low value.

4. RELATED WORK

The literature contains several technological and/or policy approaches to privacy-preserving targeting of advertisements [9–11]. Notably, Adnostic [9] provides a browser extension that encrypts the user's browsing history by means of a homomorphic cryptosystem. This allows encrypted-domain aggregation of a user's clicks, enabling an ad network to compute the amount that should be billed to various advertisers without knowing the specific advertisements accessed by the user. Our problem setting differs from Adnostic in the sense that, in addition to the user's privacy, the vendor enjoys a controlled level of privacy based on its targeting strategy. Adnostic does not explicitly compare a user's behavior with a vendor's template, though this might be possible using a suitable homomorphic encryption primitive in Adnostic's multiparty computation setting.

Another contribution that leverages multiparty computation is Privad [10], which users an untrusted dealer to anonymize a user's click behavior before providing it to the ad network. The ad network then bills the relevant advertisers and pays the publishers of the advertisements while protecting the user's privacy, and ensuring against click fraud. As in the case of Adnostic, Privad does not consider vendor-side privacy, or an explicit comparison between a user's behavior and the vendor's targeting strategy. A related scheme, called RePriv [11], provides another browser extension, which only shares the user's personal information if the user provides her explicit consent. This differs considerably

from the proposed scheme wherein the user's privacy is maintained unless she chooses to redeem her coupons.

The scheme most closely related to the proposed approach is PiCoDa [12], wherein a user can decrypt a coupon if her behavioral profile matches a vendor's target profile. Like PiCoDa, the proposed scheme allows the vendor and user to discover each other's data when the user redeems the coupon. Architecturally however, the schemes are different: PiCoDa uses Locality Sensitive Hashing [13, 14], while we use error correcting codes. Another difference is that PiCoDa does not reveal the vendor's strategy to ineligible users, whereas the proposed approach enforces a tradeoff between the information revealed to ineligible users and the targeting strategy; specific targeting reveals few bits, indiscriminate targeting reveals more bits.

5. DISCUSSION

The preceding arguments suggest that, as approximately $n - m$ bits of the vendor's strategy are revealed, m (the number of bits in the key $\mathbf{k}$) must be large enough to ensure that the vendor achieves strong privacy against brute force attacks. This implies that the behavioral vectors $\mathbf{u}, \mathbf{v}$ have to be long enough, since $n > m$. The practical implication is that the behavioral profiles must be composed of a sufficiently rich set of attributes. This is a non-trivial requirement, especially given that most behavioral targeting today is either simplistic, or is based on few behavioral attributes.

We are interested in exploring the real-world performance of such a scheme using block error correcting codes, such as Reed-Solomon [15] or BCH codes [16], for which well-studied decoding algorithms are available. Another area of interest is the exploration of interactive extensions of the fuzzy commitment paradigm, in which the user and the vendor engage in a few rounds of communication. This approach would incur a larger protocol overhead, but could allow even stronger privacy notions for one or both parties, e.g., requiring users (whether eligible or ineligible) to obliviously attempt decryption of coupons without discovering any information about the vendor's strategy.

6. REFERENCES

[1] K. Partridge and B. Begole. Activity-based advertising. In *Pervasive Advertising*. Springer-Verlag, London, UK, 2011.

[2] President's Council of Advisors on Science and Technology. Big data and privacy: A technological perspective. *Report to the President of the United States*, May 2014.

[3] J. Turow, J. King, C.J. Hoofnatle, A. Bleakley, and M. Hennessy. Americans reject tailored advertising and three activities that enable it. *Departmental Papers (ASC)*, page 137, 2009.

[4] G. Davida, Y. Frankel, and B. Matt. On enabling secure applications through off-line biometric identification. In *IEEE Symposium on Security and Privacy*, pages 148–157, Oakland, CA, May 1998.

[5] A. Juels and M. Wattenberg. A fuzzy commitment scheme. In *ACM Conf. Computer and Communications Security*, pages 28–36, Singapore, November 1999.

[6] Y. Wang, S. Rane, S. C. Draper, and P. Ishwar. A theoretical analysis of authentication, privacy and reusability across secure biometric systems. *IEEE Trans. Information Forensics and Security*, 7(6):1825–1840, December 2012.

[7] K. Sugiyama, K. Hatano, and M. Yoshikawa. Adaptive web search based on user profile constructed without any effort from users. In *Intl. Conf. on the World Wide Web*, pages 675–684, New York, NY, May 2004.

[8] M. Sudan. List decoding: Algorithms and applications. In *Theoretical Computer Science: Exploring New Frontiers of Theoretical Informatics*, pages 25–41. Springer, 2000.

[9] V. Toubiana, A. Narayanan, D. Boneh, H. Nissenbaum, and S. Barocas. Adnostic: Privacy preserving targeted advertising. In *Network and Distributed System Security Symposium*, San Diego, CA, March 2010.

[10] S. Guha, B. Cheng, and P. Francis. Privad: practical privacy in online advertising. In *USENIX Symposium on Networked Systems Design and Implementation*, Boston, MA, March 2011.

[11] M. Fredrikson and B. Livshits. Repriv: Re-envisioning in-browser privacy. In *IEEE Symposium on Security and Privacy*, Oakland, CA, May 2011.

[12] K. Partridge, M. Pathak, E. Uzun, and C. Wang. Picoda: Privacy-preserving smart coupon delivery architecture. In *Hot Topics in Privacy Enhancing Technologies*, Vigo, Spain, July 2012.

[13] A. Andoni and P. Indyk. Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions. In *47th IEEE Symposium on Foundations of Computer Science*, pages 459–468, Berkeley, CA, October 2006.

[14] M. Charikar. Similarity estimation techniques from rounding algorithms. In *34th Annual ACM Symposium on Theory of Computing*, pages 380–388, Montreal, Canada, 2002.

[15] I. Reed and G. Solomon. Polynomial codes over certain finite fields. *Society for Industrial & Applied Mathematics*, 8(2):300–304, 1960.

[16] R. Bose and D. Ray-Chaudhuri. On a class of error correcting binary group codes. *Information and control*, 3(1):68–79, 1960.