Loyal and Private:

Incorporating Customer Privacy into Loyalty Programs

6.857 Spring 2017
Oleksandr Chaykovskyy,
Pratyush More,
Zygimantas Straznickas

Introduction

Loyalty Schemes are ever present these days. From the award miles of major airlines to "get your 6th coffee for free" stamp card at your local coffeeshop, such schemes try to reward loyal customers and lure them back to spend more money. While these often sound like a win-win situation, at the core of these programs is a serious conflict. On the one hand we have retailers, who are trying to sell as many products as possible and collect as much information as possible (in most cases). On the other hand we have customers, who are increasingly concerned with their privacy.

Loyalty programs come in a variety of different forms: physical retailers, like Walmart, online retailers, like delivery.com, and even third party startups, like Stamp Me. Stamp Me advertises itself as a seamless user experience that lets customer use a variety of loyalty programs through one app. The downside of this convenience is advertised to merchants: all of the customer data is readily available [1].

Popular retailers, such as Walmart and Target are no better. As they are pushing their way into online sales, they are becoming infamous because of the amount of user data they collect. One particularly worrisome story describes Target's data mining efforts.

> *"As [Target's statistician] computers crawled through the data, he was able to identify about 25 products that, when analyzed together, allowed him to assign each shopper a "pregnancy prediction" score. More important, he could also estimate her due date to within a small window, so Target could send coupons timed to very specific stages of her pregnancy."[2]*

Armed with this data Target was able to customize their advertisement and in-mail coupons to lock-in a loyal and often rampant customer group - expecting parents.

But perhaps, with their vast data silos, they were a little too good at finding target customers. In Minneapolis an angry man stormed into a local Target and demanded explanation for why his high-school daughter was receiving coupons for diapers and baby cribs. The startled manager

profusely apologized. However, when the store called in few days later to further apologize, it turned out that the girl was indeed pregnant.

This story, however funny, puts under the spotlight the conflict between the interests of customer and merchant. In this write up we will address this conflict with a proposed loyalty preserving scheme. We will describe the goals of such scheme, overview the existing schemes and how they address these goals and propose several further improvements.

Principals

Let us first take a look at the two main principals in this system - the customer and the vendor.

Customers

We classify Customers into three groups, according to the level of their privacy concerns [3]:

1. **Privacy unconcerned**, willing to reveal their data for small incentives
2. **Pragmatic**, right rewards can convince them to reveal their data
3. **Privacy concerned**, want to protect their data, regardless of the reward

Of course this grouping implies that the customers are informed about the impact of their decision, and make educated decisions, based on that information.

Approximate distribution of user interest in loyalty programs can be estimated from the eMarketer report, which states that 80% of customers find loyalty programs worthwhile, 30% are worried that they require too much personal information, and 24% say that they would not join a loyalty program because it is a risk to their privacy [4,5].

Vendors

We classify vendors into two groups as well, based on the level of importance customer data holds to them:

1. **Data inclined:** would like to obtain customer data, and use it for their business needs, e.g. Amazon
2. **Data uninclined:** have no use for customer data, and are indifferent to having it or not, e.g. La Verde's, a small grocery store on MIT's campus

Goals

Before we can start talking about privacy preserving schemes themselves, it is important to formally establish the goals of the two principals in this system. This is because the scheme itself would be built to achieve these goals.

- **Customer-Side**
    - **Anonymity:** Provide anonymity to customers during the entire process, if so desired.
    - **Control:** Ensure control and decision-making power over private information - in terms of level of disclosure etc.
    - **Flexibility:** Enable claim of rewards based on (i) past purchases and (ii) the customer's willingness to reveal increasingly specific information. This customer should be able to make this decision as late as when s/he is claiming rewards.
    - **Simplicity:** Provide a simple customer experience and not require any additional hardware or software installation.
- **Vendor-Side**
    - **Customer Retention:** Encourage customers to return and make recurring purchases
    - **Legitimacy:** Ensure that a user's reward claim is legitimate
    - **Easy and Cheap Integration:** Should be easy to integrate with existing systems with minimal additional cost.
    - **User Data Extraction:** Encourage customers to reveal information about themselves and their purchases so that vendors can **(i) profile them / push personalised products to them** and/or **(ii) link distinct transactions and thus obtain global insights about market and demand trends**.

Adversaries

There are two types of adversaries a privacy preserving loyalty system would need to be concerned with.

The first would be an **internal** adversary. The above goals make it quite evident that the vendor and customer could act antithetically to each other, and therefore be adversaries for each other.

For example, the provider would try their best to de-anonymize a customer by using their purchase location, timing, items basket, payment method information, network metadata etc. Simultaneously, the customer would want to circumvent the system and gain illegitimate reward points.

The second is an **external** adversary - an adversary who is not one of the actors in the system. Attacks from such adversaries could include, but are not limited to, denial of service attacks and unauthorized security breaches. This second set of problems is more general, rather than specific to this particular scheme, and thus will not be addressed in this paper.

Properties

Given the above goals of the system's principals, and the adversarial threats we need to account for, the properties required of a privacy-preserving loyalty scheme are:

- **Total Anonymity:** The system should have the ability to provide total anonymity to customers throughout the entire process - during purchase, reward claim, as well as redemption.
- **Selective Revealing:** Customers should be able to reveal information about select purchases without it compromising any information about other purchases. Further, customers should be able to control how much information they wish to provide during the reward claim. This allows vendors to obtain more information about customers by giving out higher rewards in return.
- **Unlinkability:** Customer should be able to submit loyalty point to the retailer for redemption, while retailer should not be able to link this point to the purchase during which it was issued.
- **Unforgeability:** Customers should not be able to forge receipts or loyalty points.
- **Double-Spending Prevention**: Customers should not be able to redeem their valid loyalty points more than once.
- **Unbinding on Customers:** Customers should only have to decide how much information s/he wants to reveal while claiming rewards. Therefore, no binding decisions need to be made during the initial purchase itself.

In addition retailers might want to prevent customers from selling loyalty rewards. Or they might want to restrict it to exchange between family members through a family plan etc. In this case, we would need an additional property (to different extents):

- **Non-Transferability:** Vendors should be able to ensure that a particular transaction was made by the particular customer who is claiming a reward on it.

Overview of the existing schemes

We have researched a number of proposals aimed at designing valid Privacy-friendly loyalty schemes. The one we thought was closest to meeting the goals described before was ”An Advanced, Privacy-Friendly Loyalty System.”[7]

Existing schemes are based on three cryptographic primitives. First is Commitment. It allows one party to commit to particular data without revealing it. Second is Blind Signature. These can be thought of as envelopes, lined with carbon paper on the inside. As one party puts their data in the envelope and seals it, the other party cannot see what is inside. It does, however, allow the second party (authority) to blindly sign the contents. This might sounds sketchy and useless, since we wouldn’t expect an authority to sign something they do not see. And that is where the third primitive comes in: Zero Knowledge Verification (ZKV). Using ZKV protocols, authorities can verify the validity of the envelope’s contents, without learning any information the user would want to preserve.

There are three stages of the protocol we are concerned with:

- User Registration
- Earning Loyalty Points
- Spending Loyalty Points

User Registration

In order to use the loyalty scheme a user needs to register and obtain proper credentials. One way to do so is to contact a trusted issuing party. There the customer would provide all the necessary information about her identity and payment methods. In addition, the customer chooses a random number and includes it with her credentials. However, only a commitment of this random number is sent to the issuer. The issuer applies a random offset to the value and the result becomes the customer’s “loyalty secret.” This protocol keeps the loyalty secret hidden from the issuer. Furthermore, if a protocol such as a multi-show Idemix credential is used, the issuer can even be the retailer itself, since usage of credentials would be unlinkable to the registration stage[5].

Earning Loyalty Points

During the purchase, the customer puts together the number of loyalty points, timestamp and the loyalty secret. She then creates a commitment to this bundle and furthermore blinds this

commitment. The retailer then verifies its contents using ZKV. If verification passes, the retailer signs it. The customer can then remove the blind, and keep the signed commitment until she decides to trade it in.

An important security enhancement concerns the amounts used in these transactions. In order to prevent linkability these should not be processed in amounts like 39.95, but split into blocks, similar to paper money: 1, 5, 20, 100 etc.

Spending Loyalty Points

When the customer decides to trade in his loyalty points, he can present the signed commitment, along with the supporting information (number of points it's mapped to, timestamp). In order to prevent forgery and double spending it is important for the store to keep track of all the commitments it issued, as well as all of the commitments it received for trade-in. The retailer further verifies the validity of the signatures, and rewards the user with the corresponding amount of points.

If the retailer wants to prevent sharing of the points between customers, it can enforce the usage of the loyalty secret in the user's credentials. If used as a ZK proof it would prevent sharing, while keeping anonymity of the customer.

Serious Limitations

Most published privacy preserving loyalty schemes, including the one described above, are designed as cryptographic systems. Given a set of assumptions and a model of interacting with the store, they propose loyalty schemes that do preserve customers' privacy. Unfortunately, those assumptions are not entirely realistic - they either require buyers to significantly change the way they interact with the store, or underestimate the tracking power of the providers. In fact, there are a few fundamental problems that make implementing privacy preserving loyalty schemes in the real world very difficult.

Credit Cards

It is well known that debit and credit cards do not preserve their owners' privacy. Each card is uniquely associated with its owner and can be used to link individual transactions together. This means that any loyalty scheme that claims to preserve user privacy must not allow using credit cards for payment. The published schemes acknowledge this: the Blanco-Justicia et. al. paper [1] states, as an assumption, that "[t]o prevent straightforward profiling by the vendor, payment should be anonymous." The problem, of course, is that making anonymous payments in the real

[1] Blanco-Justicia et. al. Privacy-Preserving Loyalty Programs. 2015.

world is not simple. In 2016, almost 90 percent of customers who own a debit or credit card reported that they prefer to use it in stores[2]. Only 10 percent prefer using cash, the only widely supported anonymous payment method. Therefore, implementing any privacy preserving scheme in the real world will require almost 90 percent of all shoppers to change their shopping habits, which is not a simple task.

Online Shopping

In the physical world, as long as the shopper is willing to use cash, they can shop completely anonymously -- there is no way for the store to recognize them. The situation is completely different when shopping online. There are many advanced fingerprinting techniques that make it possible for website operators to identify their users. Most of these methods require no consent from the user, and often can only be noticed by using custom browser extensions. For example, in a recently published paper[3] about cross-browser fingerprinting, the researchers were able to successfully identify 99.24% of users by utilizing "many novel OS and hardware level features, such as those from graphics cards, CPU[s], and installed writing scripts." With enough effort it is possible to deceive these fingerprinting methods by using the Tor network through the Tor browser in a virtual machine. However, this is both impractical and still risky, because the use of privacy tools can be used as an identification feature by itself. Therefore, achieving transaction unlinkability online seems to be incredibly hard. Most likely, a practical privacy preserving loyalty scheme will have to disallow general-purpose browsers and only support online interactions from custom privacy-friendly interfaces.

Single Transaction Privacy

When trying to design privacy preserving schemes, it is important to understand the capabilities of the adversary. In this case, one very powerful capability of retailers is to log all transaction details. They know the time of the purchase, its location, and which cash register was used to process it. Moreover, even if the store claims to be privacy friendly and claims it does not collect this data, there is no reasonable way for the store to prove this.

Using these facts, the store can also recover a very useful piece of information - which items were bought as a part of single transaction. This has important consequences for the development of privacy friendly loyalty schemes. Specifically, it makes a certain feature, which we call Single Transaction Privacy, impossible to achieve.

---

[2] Sarah Hartman. 2016 U.S. Consumer Payment Study. 2016.

[3] Cao, Yinzhi, Song Li, and Erik Wijmans. "(Cross-)Browser Fingerprinting via OS and Hardware Level Features." Proceedings 2017 Network and Distributed System Security Symposium, 2017.

Consider a situation where a client buys 10 items from the store as part of a single transaction. Nine of these items are perfectly ordinary, but the tenth one, item X, is of a private nature. Ideally, the client would want to claim loyalty rewards for the nine items, but hide the tenth item from the store. In fact, some of the published privacy schemes specifically consider this use case. However, the information gathered by the retailer makes it impossible to hide item X in this scenario. This is because with overwhelming probability, the combination of items bought in that transaction is unique - no one else has bought this specific combination of ten items. Moreover, if the user only reveals 9 items to the store, it is still exceedingly likely that there is no other 10-item transaction that includes those nine items. Therefore, it is easy for the store to learn about item X by identifying the client's transaction from these other 9 items.

To estimate the viability of this attack, we developed a quick probabilistic model. Assume that there are 10 000 distinct items in the retailer's stock, and, since some items are more popular than others, at the probability of a person choosing to buy item x is proportional to $0.2\exp(-0.2x)$ - the exponential distribution. Then, we estimated the probability of the following event: *person A bought 10 items. What is the probability that person B, who also bought 10 items in total, bought 9 items that are the same as person A?* We found that the probability of such an event is in the order of 10^{-15}. In other words, 10^{15} transactions have to be made on average until the store can no longer uniquely identify item X.

Now, for a particular transaction, the seller would already know that this combination of 10 items were bought together, and so this limitation would not reveal any additional information about the customer's purchases. Consider a case, though, where a particular user makes multiple purchases (using the same identifier) and claims rewards for individual transactions. In such a case, the vendor would be able to learn what item X is, and connect it to the other transactions, thus gaining previously unknown information.

Therefore, if a privacy preserving loyalty card scheme allows customers to hide particular items from their transactions, it cannot allow them to claim rewards for individual transactions. Instead, the user must wait until they have made multiple transactions and claim them together, to make the previously discovered attack impractical.

Trusted Third Party System

Ultimately, we want a privacy preserving scheme that works in practice. This means the aforementioned problems are not simply interesting, but necessary to solve. It then becomes important to answer the following: How can we adapt and mould current schemes to overcome the limitations that are imposed upon us? What are the tradeoffs we make in doing so, and why are they justified?

One possibility could be the introduction of a trusted third party (TTP). Not only would TTPSs help get around the presented problems, but also provide numerous additional desirable properties. A TTP would act as the middleman between the consumers and the retailers, collecting most of the data about consumer purchases, but only revealing some of that data to the retailers, as long as it does not compromise consumer privacy.

One key property of such system is that the third party has to be trusted. Ideally, it would not be a new company developing a separate product, but a large established technology firm that already has a reputation for being trustworthy and handling user data responsibly. If such a company developed a new privacy preserving loyalty scheme product, it would be disincentivized from betraying user trust and revealing all of their data to the retailers - if it did so, all of its other products would also suffer a reputation loss.

An established company acting as a TTP also has other advantages. First, a number of these companies, including Facebook and Google, have their own payment systems. These are quickly rising in popularity, with many people using them as an alternative to credit cards. While the systems might not necessarily be private by default, TTPs can make them private if they want to. If such a TTP also required all payments to go through this private payment system, the issue of identifying customers by their credit cards would be solved. This is also true for the internet shopping problem - a TTP with an established platform could ensure that purchases happen through an interface that does not allow fingerprinting techniques to be used.

Trusted third parties can also solve the incentive problem: for retailers, using a TTP to run their loyalty schemes, even privacy preserving ones, might be more useful than running their own. This is because the same TTP will probably be used by multiple retailers, and can gather data about each customer interacting with multiple retailers. For smaller companies, this extra data the TTP provides might compensate for any data lost because of the privacy-preserving nature of the TTP and further help them obtain a competitive edge.

For the customers, using such a TTP would not feel much different from using a normal loyalty scheme. They would simply make purchases through a TTP gateway and see their loyalty points there. How retailers should be able to use the TPP is a much harder question. On one hand, they should be able to extract as much business information as possible to justify using the system. On the other hand, they should not be able to extract any information that violates a customer's privacy. One way to achieve this is by splitting retailer-side interaction into two general parts:

- **Personalization API**: retailers should be able to personalize their offers to customers based on their previous purchases. Obviously, retailers should not be able to link a customer to a physical identity because that would compromise privacy. However, they

should be given the tools to personalize the shopping experience of black-box ‘identities’ given their purchase histories.

- **Global Insights API**: retailers should also be able to use the TTP data to extract global statistical insights about their customer base, as they normally do from data gathered by loyalty programs. To still preserve customer privacy, differential privacy techniques could be used. They are especially applicable here, because only the TTP is in control of the data, so it can expose only those specific API calls that satisfy differential privacy conditions.

Conclusion

Evidently, privacy friendly loyalty schemes can be useful to both customers and retailers. In a time when public trust is eroded by data leaks, it is important for companies to reassure customers that their privacy is protected by more than just their word. However, as we also can conclude, implementing a completely cryptographic scheme, while feasible in theory, would be quite impractical in the real world. Consequently, we have to turn to more prosaic solutions, such as involving a Trusted Third Party into these transactions. Where this approach lacks in novelty and scientific finesse it gains in simplicity, flexibility, and deployability.

Sources:

[1] http://stampme.com/

[2] How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did, Kashmir Hill, Forbes, Feb. 16 2012. https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#63c11ac76668

[3] Oliver Hinz, Eva Gerstmeier, Omid Tafreschi, Matthias Enzmann, and Markus Schneider. Customer loyalty programs and privacy concerns. In *Proceedings of BLED 2007,* 2007

[4] http://www.businessnewsdaily.com/4615-loyalty-programs-privacy-concerns.html

[5] https://www.emarketer.com/Article/Keep-Users-Happy-Loyalty-Programs-Must-Walk-Fine-Line/1009958

[6] Blanco-Justicia et. al. Privacy-Preserving Loyalty Programs. 2015.

[7] Milutinovic M., Dacosta I., Put A., De Decker B. (2014) An Advanced, Privacy-Friendly Loyalty System. In: Hansen M., Hoepman JH., Leenes R., Whitehouse D. (eds) Privacy and Identity Management for Emerging Services and Technologies. Privacy and Identity 2013. IFIP Advances in Information and Communication Technology, vol 421. Springer, Berlin, Heidelberg