Analysis of Tokenization in Digital Payments

By Flora Liu

Mentor: Ming Chow

Cyber Security Fall 2016 Final Paper

Abstract

Payment security remains a challenge for online platforms that handle sensitive information, such as credit and debit card numbers. Though innovative solutions as well as stricter industry standards have been developed to tackle the issue of security breaches, cyberattacks targeting online payment systems continue to happen regularly. This paper will focus on tokenization, a method applied to protect sensitive data in payment transaction. The way in which tokenization, combined with encryption, enhances security will be dissected in the discussion below. In addition, I will also explore applications of the token system and uncover possible weaknesses in its implementation.

Intro

With the click of a few buttons, one can purchase a whole new wardrobe, order a camera drone, or buy groceries without leaving the comfort of home. Millions of online payments take place every day on e-commerce platforms, and this number is projected to grow.[10] According to Visa, e-commerce sales are expected to reach over $620 billion by 2018.[11] These digital payments deliver the promise of smooth transactions. However, online payment processes handle highly sensitive personal data, and as a result, have turned these systems into battlefields for cyberattacks.

When a consumer makes an online payment, the cardholder's information (PAN short for Primary Account Number) is entered into the merchant system and sent to a payment processor.[12] The processor, sometimes a third party, is responsible for passing the card data to a payment brand (such as Visa, MasterCard, American Express, etc.), which then requests for a verification from the issuing bank. If the card data is legitimate, the bank will return an authorization number along with the PAN to the payment brand.[12] The payment brand will forward this authorization code back to the processor, which sends it to the merchant who completes the sale. Unfortunately, cyber criminals are becoming more sophisticated when attacking payment systems to attain credit card data. They maliciously target each step of the payment processing chain described above to commit fraud or identity theft. Illicit methods include sniffing or diverting the data to a different destination. Let's explore how tokenization can increase data security and reduce vulnerabilities to attackers.

To the Community

The expansion of Internet access to people all over the world has created a global marketplace, one in which transactions transcend national borders, time zones, and industries. E-commerce has not only transformed the traditional way of selling and buying, but also led to skyrocketing transaction volumes and new payment systems. This proliferation of technology provides many benefits to merchants and consumers, but how the digital process works under the covers remains an enigma for many. Users may not have the technical knowledge or background to understand fully how the online payment chain works. People use technology even though they are not aware of the potential privacy and security risks. Since consumers mainly interact with abstracted interfaces, they assume that security measures have been put in place to protect their personal data. Thus, conversations about security challenges and flaws of online payment systems are limited. The intention of this paper is to enlighten and engage the community in the issue of privacy concerns behind digital transactions. I hope to inform the audience about the protection of sensitive payment information by way of tokenization. I challenge them to raise questions regarding their own use of digital payment technologies and educate others about cybersecurity.

What is Tokenization?

Tokenization is "the process of protecting sensitive data by replacing it with an algorithmically generated number called a token".[9] In the context of digital transactions, this tokenized surrogate of meaningless letters and numbers is used in place of the original cardholder data.[6] Tokens "can even have the exact format (size and structure) as the original value", which is useful for minimizing any application changes or updates.[8] Once a payment transaction is authorized by the issuing bank, the original sensitive data is stored in a highly secure, offsite server or "vault". This location contains a reference database where the token can be exchanged for the real account data. Meanwhile, the token is passed along the online payment processing chain to accomplish the necessary tasks without exposing the real PAN. Unless one has access to the original key used to create the token, the tokenized result is not mathematically reversible.[9] Even if an organization's security was breached, the tokenized data does not contain real card information, so it has no value to external parties outside the native application environment.

Tokenization certainly adds a layer of protection to the transit of sensitive data by limiting exposure between various servers, programs, and parties (merchants, processors, acquirers, and issuers). It reduces handling of and access points to the true card information. Tokenization removes sensitive data from the merchant environment as quickly as possible (root cause of data security), eliminating any need to store actual credit card numbers. The use of tokens doesn't interfere with the frictionless shopping experience of consumers either.

Encryption and Tokenization

Cardholder data should never be sent in the clear without any encryption. Otherwise, it is easily readable and can be exploited by anyone who has access to it. Encryption is "the transformation of data into a form unreadable by anyone without a secret decryption key". [7] Encryption converts sensitive data to cipher text at the origin and requires decryption to get back the original form at the destination. For a given input, encryption will consistently produce the same result. Tokenization, on the other hand, replaces sensitive data with an algorithmically transformed placeholder. A key difference between tokenization and encryption is that encryption is reversible.

A point to note is that tokenization is sometimes mistakenly seen as a "more effective" way to protect data than encryption. I would argue that in fact they provide benefits in different aspects of data protection, and are best used together as a dual solution. Both tokenization and end-to-end encryption are put to practice for a common purpose: limit systems from accessing the actual credit card number and masking sensitive information. An effective method to maximizing payment security therefore entails first encrypting then tokenizing. [2] At the point of capture, the PAN should immediately be encrypted for protection so that the information is unreadable during transit activity. If authorization is then issued by the bank, a randomly generated token value should be given back to the merchant's system for use. Encryption combined with tokenization can protect cardholder data in transit and eliminate possibilities of exploitation. Hence, the layered effects of tokenization and encryption enhance the security and integrity of sensitive information.

Industry Standards

The PCI (Payment Card Industries) established a set of security standards to ensure effective protection of payment card information. [5] This framework known as the PCI Data Security Standards (PCI DSS) guides merchants on how to secure sensitive data. PCI DSS is especially controversial to businesses, due to the complex process and expensive costs of compliance. [4] They recognize the importance and need for data security, but view payment processing as an additional obstacle. Thus, digital payment processing is often outsourced to third parties to avoid merchant contact with cardholder data.[4]

Examples of data protection methods in the DSS suggested include: one-way hashes, truncation, index tokens and pads, strong cryptographic algorithms, as well as encrypted transmission of cardholder data.[3] PCI DSS provides recommendations for implementations of tokenization and guidance on tokenization as a solution to data security. However, the lack of general standards for tokenization practices remains a challenge due to possible confusing or conflicting implementations.

Apple Pay, the new payment processing mechanism introduced by Apple, makes use of tokenization.[9] After a user inputs his or her credit card information into Apple Pay, Apple sends the details to the bank. If the cardholder data is validated, the bank will generate a token and give it back to Apple. The token is used in place of the credit card information for purchases. Thus, sensitive user card information is stored securely via tokenization in Apple phones and these randomized token numbers are worthless to attackers.

Tokenization Vulnerabilities and Criticisms

In 2010, Visa published standards for tokenization practices. To generate tokens, the use of a known strong cryptographic algorithm or one-way irreversible function was suggested. For multi-use tokens, Visa recommended hashing the data with a unique salt value per merchant. Nevertheless, using the same salt for the same merchant is a dangerous practice. If a malicious individual is able to retrieve the salt and proceed to conduct a dictionary attack, the payment card information could potentially be recovered. [3] To overcome this problem, issuers should create multi-use tokens from a pre-generated token table so that there is no way a token can be reverse engineered.[4] A random token should be selected so that there is no specific relationship between the PAN and the token.

In 2014, Visa launched the "Visa Token Service". Visa's tokenization efforts were mostly lauded but met with some criticism regarding security vulnerabilities. For example, cyberattack targets could shift to tokens databases (where the relational maps between tokens and PANs are kept). The tokenized architecture of payment mechanism may reduce security flaws on the merchant side, but could transfer security risks upstream to the acquirer or issuer. Furthermore, access to centralized or cloud tokenization systems bring up problems of latency and resilience. [1] Token generation and processing involves additional cost in the online payment processing chain. Other concerns regarding tokenization include network segmentation (separating tokenization and data processing system) and uses of recycled tokens.

Conclusion

Tokenization of sensitive cardholder information in online payment systems reduces security risks significantly. Because tokens are generated by strong cryptographic algorithms, they are difficult to reverse engineer and decode. The major security benefits gained from token usage are worth the additional effort of adding tokenization in the online payment processing chain. Tokenization puts to practice the principle of least privilege by limiting cardholder data exposure to selected systems. Layering end to end encryption with tokenization will preserve the integrity of data and allows digital transaction platforms to stand a better chance against malicious behavior. Users making online purchases should be mindful of the sensitive data they are releasing, while merchants and banks need to build payment systems with security considerations from the start. As the trend of digital transactions continues to grow, cybersecurity of cardholder data is an issue that will affect more users. Tokenization is among one of the most cost-effective methods out there to balance the security needs of consumers, businesses, and banks.

References

1. Carrol, Pat. "Tokenization: 6 Reasons The Card Industry Should Be Wary". http://www.darkreading.com/perimeter/tokenization-6-reasons-the-card-industry-should-be-wary-/a/d-id/1316376
2. First Data. "EMV and Encryption + Tokenization: A Layered Approach to Security". https://www.firstdata.com/downloads/thought-leadership/EMV-Encrypt-Tokenization-WP.PDF
3. Kuo, Li-Hsiang. "Cracking Credit Card Number Tokenization". http://pages.cs.wisc.edu/~lorderic/webpage/tokenization-crack.pdf
4. Mercator Advisory Group. "Closing the Gap in Tokenization: Removing the Last Vulnerability". https://www.voltage.com/wp-content/uploads/Mercator_Closing_the_Gap_in_Tokenization_Removing_the_Last_Vulnerability.pdf
5. PCI Security Standards. "Tokenization Product Security Guidelines". https://www.pcisecuritystandards.org/documents/Tokenization_Product_Security_Guidelines.pdf
6. PCI Security Standards. "Tokenization Guidelines Info Supplement". https://www.pcisecuritystandards.org/documents/Tokenization_Guidelines_Info_Supplement.pdf
7. Roberts, Eric. "Encryption F.A.Q." https://cs.stanford.edu/people/eroberts/courses/cs181/projects/1995-96/clipper-chip/encyptfaq.html
8. Securosis. "Understanding and Selecting a Tokenization Solution". https://securosis.com/assets/library/reports/Securosis_Understanding_Tokenization_V.1_.0_.pdf

9. Squareup. “What Does Tokenization Actually Mean”. https://squareup.com/townsquare/what-does-tokenization-actually-mean/
10. Statista. “Number of Digital Buyers Worldwide From 2014-2019”. https://www.statista.com/statistics/251666/number-of-digital-buyers-worldwide/
11. Visa. “Visa Digital Solutions”. https://usa.visa.com/partner-with-us/payment-technology/visa-digital-solutions.html
12. Visa. “Security Tokenization Infographic”. https://usa.visa.com/dam/VCOM/Media%20Kits/PDF/visa-security-tokenization-infographic.pdf