PROTECTING YOUR SHOPPING PREFERENCE WITH DIFFERENTIAL PRIVACY

N.SRINIVASA RAO [1],LAVANYA K[2].

[1] Assistant Professor, **DEPT OF MCA, SKBR PG COLLEGE , AMALAPURAM, Andhra Pradesh**

Email:- naagaasrinu@gmail.com

[2]**PG Student of MCA, SKBR PG COLLEGE , AMALAPURAM, Andhra Pradesh**

Email:- lavanyak266@gmail.com.

ABSTRACT— Publishing datasets plays an essential role in open data research and promoting transparency of government agencises. However, such data publication might reveal users' private information. One of the most sensitive sources of data is spatiotemporal trajectory datasets. Unfortunately, merely removing unique identifiers cannot preserve the privacy of users. Adversaries may know parts of the trajectories or be able to link the published dataset to other sources for the purpose of user identification. Therefore, it is crucial to apply privacy preserving techniques before the publication of spatiotemporal trajectory datasets. In this project, we propose a robust framework for the anonymization of spatiotemporal trajectory datasets termed as machine learning based anonymization (MLA). By introducing a new formulation of the problem, we are able to apply machine learning algorithms for clustering the trajectories and propose to use k-means algorithm for this purpose. A variation of k-means algorithm is also proposed to preserve the privacy in overly sensitive datasets. Moreover, we improve the alignment process by considering multiple sequence alignment as part of the MLA. The framework and all the proposed algorithms are applied to T-Drive, Geolife, and Gowalla location datasets.

Index terms— Machine Learning, Protecting Transaction, Online Shopping, Payment.

I. INTRODUCTION

In the last decade, online banks were commonly used to provide financial services [1]. However, online banks are vulnerable to outsider [2] [3] and insider attacks [4] [5]. Outsider attacks include brute-force attacks [6], distributed attacks [7] and social phishing [8]. Insider attacks are data misused by people with authorized access.

Outsider and insider attackers can collect the financial information of consumers to infer personal shopping preferences, consumption patterns or credit statistics [9] [10]. If consumers' shopping records are disclosed, consumers may receive advertisement recommendation, harassing message and fraud emails. More seriously, it contributes to loan promotion, illegal investigation, property fraud, and even kidnapping [11]. If consumers have no reasonable assurance of their accounts, they would be reluctant to use online banks, leading to user loss and higher cost for online banks. Therefore, appropriate methods are required to stem the erosion of privacy rights in online banks.

To protect consumers' privacy, existing approaches mostly used cryptography. Cryptography schemes mainly utilized encryption technology [12] [13] and authentication technology [14] [15], which could prevent illegitimate and unauthorized access. However, it is generally difficult for cryptography schemes to handle insider attacks effectively. Insider attackers can still misuse their authorized access to obtain credit statistics and shopping records [16].

On the other hand, differential privacy can provide strong privacy protection by ensuring the indistinguishability of one entity involvement in the dataset [17]. However, directly applying differential privacy in online banks incurs some problems.

To address these challenges, we propose an optimized differential private online transaction scheme (O-DIOR), in which we define a new noise probability density function. The fundamental strategy is to basically eliminate the probability that noise is generated beyond the boundaries. The scheme can satisfy the differential privacy definition because the noise can be any value in a valid range to avoid the case that the consumption amount and noise can be inferred.

Considering the consumption amount may be great and there is not enough money to generate the noise, we propose a revised O-DIOR scheme (RO-DIOR) to select variable boundaries. We define a new parameter in the noise distribution to adjust boundaries at a time point. We adjust the noise distribution to increase the probability of saving money from a payment application when the consumption amount approaches to zero and increase the probability of withdrawing money from the payment application when the consumption amount approaches to maximum.

To implement the scheme, we design a security module for an online payment application to generate and eliminate the noise to guarantee the utility of consumption amounts. Here we take Apple Pay for example. In our scheme, a consumer uses Apple Pay to pay for his bill, obtaining money from his online bank account and Apple Pay account. Apple Pay does not store consumers' card numbers and consumption records that can track consumers, so it cannot know consumers' shopping preferences. Traditionally, Apple Pay directly withdraws money from online banks, our additional step is to use money from consumers' own Apple Pay accounts, which may not incur more security and trust problems.

II. LITERATURE SURVEY

Online banks have been commonly used for payment services. A great deal of work aims to protect the online consumption privacy for higher privacy-level performance. The approaches can be classified into two categories. The first category is authentication. This work in [20] first described a systematic multi-factor biometric fingerprint authentication approach which provided an identity verification process for validating the legitimacy of remote users. They developed a privacy protection gateway for obscuring and desensitizing the consumers account details using tokenization and data anonymization techniques. The study in [2] showed that the authentication of many Norwegian online banking consumers was too weak, and discussed authentication methods and possible attacks. The work in [21] studied authentication issues of client and transaction for online banks. Paper [22] focused on evaluating authentication methods which were used in online banks. The work in [14] used a short-time password solution and a certificate-based solution to resist the online channel- breaking attacks.

The second category is encryption. Pathak et al. [12] designed a protocol for privacy preserving bank computations using arithmetic cryptography. The work in [23] presented a secure hybrid architecture model for internet banks using Hyperelliptic curve cryptosystem and Hash algorithm. Tebaa et al. [24] proposed a hybrid homomorphic encryption method for protecting the privacy of banking data in the cloud.

However, these schemes still have some limitations. It is difficult for authentication and encryption methods to handle insider attacks in online banks, because consumption records have to be exposed to people with authorized access.

To handle insider attacks, differential privacy is widely used. To the best of our knowledge, our scheme is the first to meet the requirement of differential privacy for online banks. We compare with existing schemes which address noise boundary problems of differential privacy in other scenarios. Duchi and Jordan [18] used lower and upper boundaries for estimation of population quantities and variational boundaries on mutual information under local privacy.

Zhang et al. [25] proposed differential privacy-preserving schemes for smart meters, limiting the range of noise and capability of batteries.

Hardt and Talwar [26] gave polynomial time computable upper and lower boundaries on noise complexity and error. The work in [27] presented privacy buckets for computing upper and lower boundaries for approximate differential privacy after rfold composition. The paper [28] preserved privacy of individual entries with constrained additive noise and its optimal probability density function could maximize the measure of privacy.

III. PROPOSED SYSTEM

The overview of our proposed system is shown in the below figure.

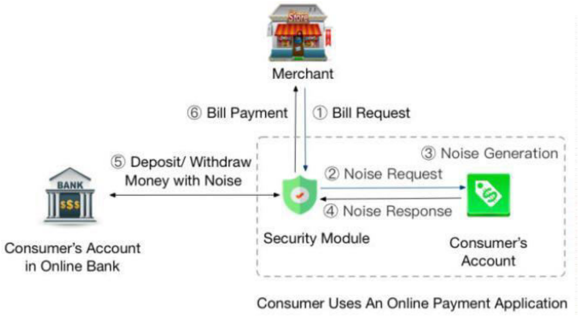


Fig. 1: System Overview

Implementation Modules

Consumer

- In this module, consumer register and login to the system. He/she can view his profile, search the product information.
- He may purchase product and pay through online payment application.
- In this he/she can view the products details, purchased product details, and payments details.

Bank Admin

- In this module, the bank admin login to system and view consumer purchases details.
- In this admin can view the consumers details, consumers payment request and he can accept or rejected.
- He can view the all transactional details.

Merchant

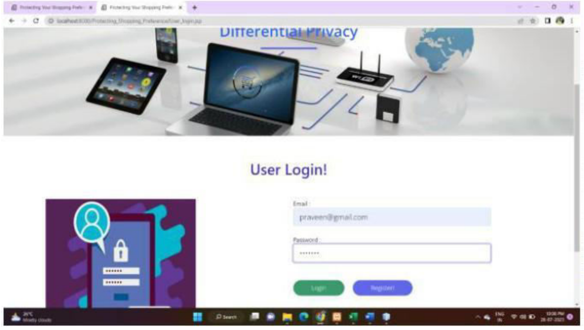
- In this module, the merchant can register and login to system. He can add product information.
- In this, merchant can view the purchase request of consumer and generate bill and send bill request to bank admin.

He can also view the payment details, all purchased products information

Payment application

The payment application could be like Apple Pay, Alipay, Paypal or Wechat pay on the mobile. It is like a money pool which can store a certain amount for a consumer. It can facilitate us to generate and eliminate the noise for the consumption amount. In this project we develop it in local host only not in real time

IV. RESULTS

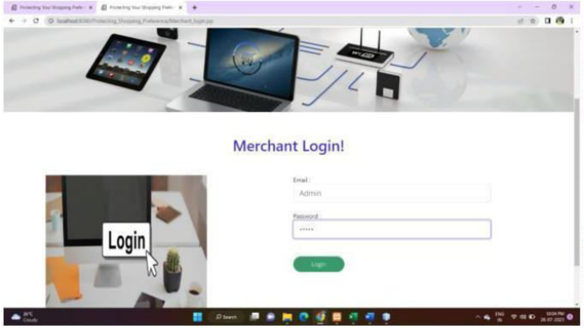


User Login!

Email : praveen@gmail.com

Password : •••••••

Fig. 2: User Login Page

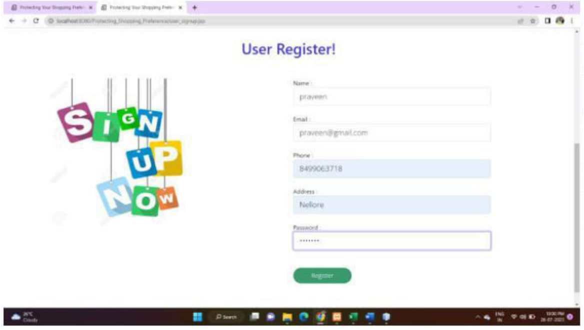


Merchant Login!

Email : Admin

Password : •••••

Fig. 3: Merchant Login



User Register!

Name : praveen

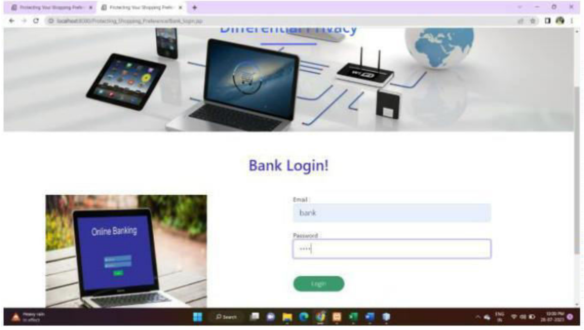
Email : praveen@gmail.com

Phone : 8499063718

Address : Nellore

Password : •••••••

Fig. 4: User Registration



Bank Login!

Email : bank

Password : ••••

Fig. 5: Bank Login

V. CONCLUSION

Protecting user data with differential privacy is a challengingproblem for online banks. The method of directly applying differential privacy is illustrated in a DIOR scheme. In this project,we propose O-DIOR, a differential private online transactionscheme to address privacy concerns during financial transactions.O-DIOR can set boundaries of consumption amountwith added noise, considering the range of account balancein reality. With a payment application as a noise generator,activities and behaviors of consumers cannot be inferred fromconsumption records. Next, we further revise O-DIOR topropose RO-DIOR, satisfying the need of selecting differentboundaries. Moreover, in-depth theoretical analysis has provedour schemes can satisfy the constraint of differential privacy.Experimental results illustrate that the relevance between thereal consumption amount and online bank transaction amountis reduced significantly, and the privacy losses are less than0.5 in terms of mutual information.

REFERENCES

[1] S. Nilakanta and K. Scheibe, “The digital personal and trust bank: Aprivacy management framework,” Journal of Information Privacy andSecurity, vol. 1, no. 4, pp. 3–21, 2005.

[2] K. J. Hole, V. Moen, and T. Tjostheim, “Case study: Online bankingsecurity,” IEEE Security & Privacy, vol. 4, no. 2, pp. 14–20, 2006.

[3] A. Rawat, S. Sharma, and R. Sushil, “Vanet: Security attacks and itspossible solutions,” Journal of Information and Operations Management,vol. 3, no. 1, p. 301, 2012.

[4] M. B. Salem, S. Hershkop, and S. J. Stolfo, “A survey of insider attackdetection research,” Insider Attack and Cyber Security, pp. 69–90, 2008.

[5] E. E. Schultz, “A framework for understanding and predicting insiderattacks, Computers & Security, vol. 21, no. 6, pp. 526–531, 2002.

[6] C. Herley and D. Florˆencio, “Protecting financial institutions frombrute-force attacks,” in Proc. IFIP International Information SecurityConference, 2008.

[7] A. Householder, K. Houle, and C. Dougherty, “Computer attack trendschallenge internet security,” Computer, vol. 35, no. 4, pp. 5–7, 2002.

[8] T. N. Jagatic, N. A. Johnson, M. Jakobsson, and F. Menczer, “Socialphishing,” Communications of the ACM, vol. 50, no. 10, pp. 94–100,2007.

[9] Y.-A. De Montjoye, L. Radaelli, V. K. Singh et al., “Unique in theshopping

mall: On the reidentifiability of credit card metadata,” Science,vol. 347, no. 6221, pp. 536–539, 2015.

[10] C. Krumme, A. Llorente, M. Cebrian, E. Moro et al., “The predictabilityof consumer visitation patterns,” Scientific reports, vol. 3, p. 1645, 2013.

[11] H. Wang, M. K. O. Lee, and C. Wang, “ Consumer privacy concernsabout internet marketing,” Commununications of the ACM, vol. 41, no. 3,pp. 63–70, 1998.

[12] R. Pathak, S. Joshi, and D. Mishra, “A novel protocol for privacypreserving banking computations using arithmetic cryptography,” inProc. Security and Identity Management, 2009.