Survey Paper on Generating Customer Relationship Management Efficiently using Homomorphic Encryption and Data Packing

Tejashree Malshikare
Computer Engineering
Dattakala Faculty Of Engineering
Swami Chincholi, Daund, Pune
tejashri.malshikare@gmail.com

Prof. Amrit Priyadarshi
Computer Engineering
Dattakala Faculty Of Engineering
Swami Chincholi, Daund, Pune
amritpriyadarshi@gmail.com

***Abstract*—** In recommender systems, recommendations are generated based on the data collected from the user. The important requirement of the basic Information Filtering architectures is to protect the privacy of all the users. By using the Homomorphic encryption and data packing the recommender system provides good privacy of customer data. The data protection system gives security from malicious third parties, but does not provide security from the service provider. In this paper, our aim is to generate the dynamic recommendations and protect the confidential data of user against the service provider while protecting the functionality of the system. This system is very useful to generate dynamic recommendations by preserving the privacy of the users.

Keywords- *Homomorphic encryption, privacy, recommender systems, Privacy Service Provider, Customer Relational Management(CRM)*

*****

I. INTRODUCTION

There are a lot of various data mining techniques available today that can be used for obtaining distribution information for any kind of data under use. Many of these techniques have limitations such as some techniques fail to provide post processing capabilities, whereas on the other hand some techniques only provide visualization feature and fail to suggest any direct actions that can improve the decision making with cost improvement.

The image database size is very large also its deployment in applications is critical task in today's technology era. Online recommender systems provide personalized services to users. The data provided by user may be misused by the service provider therefore it is very important to protect the users sensitive data so the proposed system encrypt the data and then generate the recommendations. In recommender system, the user profiles are accepted by the online services and generate recommendations by using the accepted profiles. It is very important to protect the privacy of all participants. The data protection systems gives security against third parties, but does not provide protection from service providers. This is the privacy risk for users. In this paper, our aim is to generate dynamic recommendations and provide security to the users confidential data from the service provider and protect the functionality of system. The protection of users privacy is main aim of basic information architecture. To provide the protection from malicious users the recommendation system mostly focus on the access controls and secure transmission of the user data but can't provide privacy from the service provider.

We are going to recommend encrypted private data and will process encrypted information to generating the recommendations. So, we are going to construct an efficient system that does not require the active participation of the user by using a semi trusted third party. The existing private recommendation system consists of Paillier encryption system but system is more difficult and inefficient. We solve this problem by using EIGamal algorithm. The techniques for generating recommendations for users depend on the information gathered from the user. Collaborative filtering algorithm is collect data from different resources such as users' profiles and its behaviors.

We take example as online shopping, we have to find the services and the products which are appropriate for users, so we use the click logs and preferences in the past for finding the appropriate services for users. As this increases the possibility of products purchased by user. It is useful for recommendation system.

The recommender systems requires improvements to make recommendation methods more Efficient and applicable to real-life applications, including recommending vacations, various types of financial services to investors, and products to purchase. This improvement as more efficient methods for representing user behavior and the information about the items to be recommended and also advanced recommendation method are used.[1]

A. Motivation

There are lots of problems in industries one major issues is the customer management in the industry. Our motivation is from the telecom industry and we know that there are various organizations exist and there is a hard competition among them to attract the customers towards them and also this is a fact that customers got lot of alternative options available so they keep switching from one service to another, so there is increased challenge for organization to keep clients attached to increase the organizations profit. This process is also called as customer churning; therefore our aim is to extract useful knowledge from a large customer databases and come to decision in order to overcome this customer churning. To illustrate our motivation we consider customer relationship management CRM thus we take example of the telecommunications industry. This industry is has more and more competitions in recent days The issue is that their customers switching to other industry's services. This increases industry loss. As each customer have number of

choices in telecommunications and financial services. As this customers are switching from one service provider to another. This phenomenon is called customer "churning". The customer churning is a major problem for these companies and makes it hard for them to stay profitable. It is noted that the number of such databases keeps growing rapidly because of the availability of powerful and affordable database systems.

B. Background

There are a lot of various data mining techniques available today that can be used for obtaining distribution information for any kind of data under use. Many of these techniques have limitations such as some techniques fail to provide post processing capabilities, whereas on the other hand some techniques only provide visualization feature and fail to suggest any direct actions that can improve the decision making with cost improvement. Lot of these schemes require manual work from humans in order to improve the overall net gross, therefore our plan is to develop a system that automate this process and prove to be efficient solution for post processing techniques.

II. PROPOSED WORK

Our main aim in this research work is to focus on customer churning process and develop some strong techniques to overcome the same using data mining techniques. Our algorithm does not depend only on the prediction, but also on the classification and probability estimation. User information is available from decision trees. Also we have to protect the confidential data of user against the service provider. While protecting the functionality of this system is efficient to generate dynamic recommendations in a privacy-preserving manner.

A. Features

Decision trees are used to make predictions by means of classification process; therefore they are called as predictive models. The model is generally represented in the form of an upside tree starting with the root node at the top and followed by leaf's at the bottom.

All the branches of a decision tree represent the rules; These rules can be used to obtain records falling under a predefined category. In applications having customer relationship management, decision trees classify existing customer records in to segments that has certain behavior. The process initiates with customer data related to certain behavior, in our case churned customers i.e. customers who have left the services for a competitor, and those who haven't. Decision trees are then constructed from these data, which gives the splitting attributes and the criteria which divides the customers in to two categories loyal and unloyal customers in our paper. Since the rules for classes under which customers will fall are defined, we can use these rules to classify existing customers and also predict their behavior in future. For example if a customers whose record attributes match with that of churn customers record, then there is possibility that the current customer may be also churn in some time. So this prediction is most important for marketers in order to plan activities to avoid churning. In private recommendation the privacy sensitive data such as user preferences and similarity values between users were to be encrypted and generate recommendation by processing those data. As the Homomorphic property permit us to realize linear operations in the encrypted data. Efficiency plays an important role in the success of cryptographic protocols.

B. Constraint

In online services such as electronic retailers, e-commerce and product providers always provide a large number of products which users might enforced to choose from them. The most important challenging process in online services is to matching consumers with appropriate products and that helps consumers in decision making process. It is helpful for recommendation system. A proposed recommendation for products that identify a user's choice can not only enhance user satisfaction, but also increase conversions and profits for electronic retailers. Internet leaders are highly using product recommendation engine for their recommendation, for example Amazon, Google, Netflix, TiVo and Yahoo etc. Recommender systems provide extensive technology used to promote cross-selling. Collaborative filtering method is mostly used for user recommendations. Most collaborative filtering methods require explicit user feedback, such as ratings, it is fact that users rate only a small portion of available products. As a result, the rating systems have insufficient consumer feedback which leading to failure for some recommendations.

C. Research Methodology

As mentioned earlier about the telecom industry, and the customer churning problem, basically here we will mention first about what is customer churning. As there are lots of telecom industries available people also have got a lot of choices to switch from one service to other service, this switching of customers is called as customer churning. Due to this churning problem lots of telecom industries are facing financial loss, therefore there is lot of demand from these companies for a system that can overcome the churning by making the decision making process more easy and simple.

Our aim in this research work is to overcome the churning problem by extracting actions of consumer from large data base of customers the extraction process is based on post processing techniques. When considering the datasets of customers, datasets are found to be sensitive and unbalanced, we have to work very carefully with the customer data as there is possibility that the actual customers which stay may be larger as compared to the customers who will actually churn.

Our plan is to develop a customer relationship management system with smart shopping facility, where we divide the customers in to two categories known as loyal and unloyal using decision tree algorithm. We then provide actions that one needs to take in order to convert unloyal customers to loyal. We maintain the information of customers in database, based on customers attributes. The decision tree in our project is based on important attributes of customers such as name, age, gender, birthday etc. and also considering some other information related to finance such as annual income, life style etc. One main reason behind using decision trees is that it can be converted to rules if one wishes to obtain the representation

in order to classify the results depending on certain condition. In our case we obtain the characteristics of customers using decision trees and later classify the customers in two categories either loyal or unloyal. The output of decision trees will be given as input to post processing techniques, depending on probability and prediction. For example let us consider the following figure which shows the categorization of customers on the basis of services used by them and the cost spent on that service. The techniques for generating recommendations for users depend on the way personal information of user is collected. This personal information can be provided by the user in profiles or the service provider can observe users actions as click logs. Also, more number of users information helps the system to improve the accuracy of the recommendations. The personal information of the users has a several privacy risk since there is no guarantee of the service provider that it will not misuse the user's data. It is highly seen that whenever a user enters the system, the service provider claims the ownership of the information provided by the user and authorizes itself to distribute the data to third parties for its own benefits.

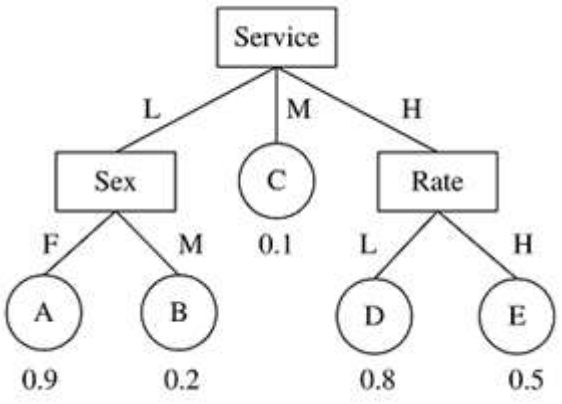

Fig.1 An Example Of Customer Profile

Once we have obtained the decision tree by using customer profile we can now make use of the decision tree to obtain the classification based on probability for e.g. probability of being loyal or unloyal. There are many data mining algorithms available, but the common problem that many of algorithms have that is they just restrict themselves to developing conceptual models. In applications like CRM, where we are not only interested in models but also expect interpretations from these models to maximize the profit. Moreover people feel satisfied with the obtained models and don't show interest in further information processing so as to maximize the profit. Also many of the existing data mining algorithms focus on developing customer profiles in order to forecast customer's characteristics belonging to certain classes. For instance what kind of customers, who all are likely to be attributors? Etc. all these profiles are based on different customer attributes. Considering from the enterprise point of view these profiles are useful. Therefore in order to make the customer relationship better, there must be some mechanism that will easily guide the organizations on what actions they need to take to change the status of customers from one state to other, considering our application of the telecom industry, the actions that can be taken to increase the profit of particular company is providing group of customers with different benefits various beneficial plans like increasing the monthly rentals, or enhancing the current services to the next level. But the only problem that come here is in order to take decision one must also consider the cost and how much particular action will benefit to the organization, at the same time it is also possible that there are more no. of actions for a particular customer profile, therefore the responsibility of choosing the best suitable action which will maximize the profits largely also comes into consideration.

There is a approach to privacy preserving recommender system is to use encryption and multi party computation techniques for users sensitive queries. There is need to encrypt the privacy sensitive data and then generate recommendations by processing under encryption.[8]

To address the privacy considerations in recommender systems, where the private user data is encrypted and recommendations are generated by applying an iterative procedure based on the ElGamal algorithm. The algorithm computes the secured information of users in a subspace and generates recommendations by calculating projections in the encrypted domain. This EIGamal algorithm takes many iterations for convergence and in each iteration users need to participate in decryption procedure, where the users are considered to being online and honest. We propose a method to protect the privacy of users based on a probabilistic factor analysis.

III. SYSTEM ARCHITECTURE

To generate recommendations, User must give two inputs: the heavily rated vector to compute the similarity values between users, and the slightly rated vector to generate recommendations of the average rating of the top most similar users. These vectors are very privacy-sensitive and thus, the service provider store this vectors in the encrypted form. The service provider does not have the decryption key, as this service provider can't access user's private data.

To generate recommendations, the service provider and the privacy service provider run a cryptographic protocol without interaction of the users. Recommendations can be generated in a privacy-preserving way during the idle time of the service provider and the privacy service provider before any user asks for recommendations. This shows that a user will get recommendations immediately after his request with no any delays.

Our research work uses the collaborative filtering. In the privacy preserving collaborative filtering users control their own private data and more number of users can compute a aggregation of their data without using individual users' data.[2]
The recommendation has two modes namely static and dynamic recommendation:

1. Static recommendation: In the static recommendation, it will generate according to the product only without considering purchase of users. Consider user 1

 ISSN: 2321-8169
6432 - 6436

purchased the software A then system recommended product B for the product A. Its work is static.

2. Dynamic recommendation: In dynamic recommendation the system will generate the recommendation according to products purchased by users. In this case consider five users purchased the product A. Out of five three will purchase product B along with product A and remaining two will purchased product C. So on the basis of maximum frequency, the system will recommend the product B for the product A for the new users

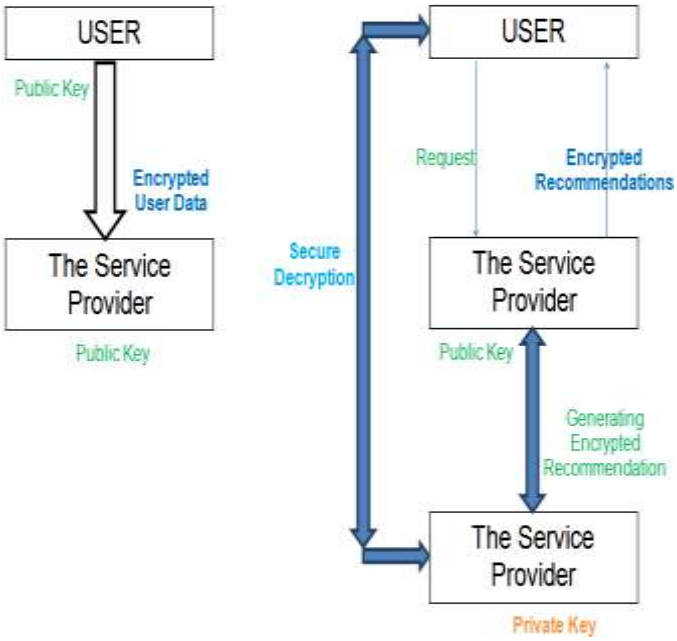

Fig.2 System Architecture

IV. SYSTEM MODEL

The design of proposed system is to be implemented to execute using computer. Hence design should be converted to the computer understanding language. Implementation is a stage in the project to convert the physical system specifications into working and reliable software .It is the stage of a project during which theory is turned into practice. It is very critical and most important step in achieving a successful new system. So it provides user with guarantee that developed system will work effectively.

In CRM, we focus on the output of decision tree algorithms as the input to our post processing algorithms. Our algorithm gives not only a security, but also a probability as the classification, such as the probability of being loyal. As Current systems require active participation of user which has a privacy risk. To overcome this problem we eliminate the need for active participation of users while using a semi trusted third party that is the Privacy Service Provider (PSP), which is allowed to perform the assigned tasks properly, but is not allowed to see the private data of user. Encryption and Decryption are doing using Homomorphic encryption algorithm such as ElGamal algorithm. Using this PSP users upload their encrypted data to the service provider and by using a collaborative filtering technique between the service provider and the PSP, the recommendations are generated without users active participation.

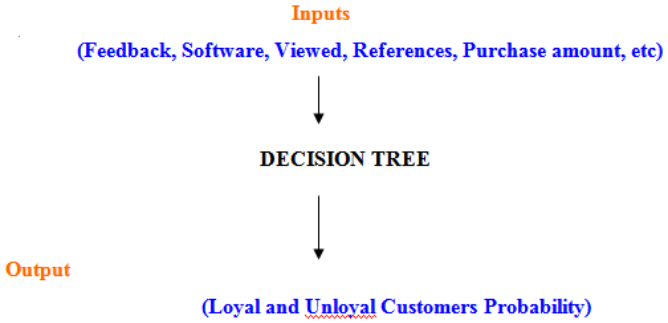

Fig.3 System Model

Suppose that for a customer A, the attribute Atr has an original value X. To change its value to Y, an action is needed.

X is probability that we got…
Y is what we expecting..!!
Therefore Action U is to be taken on customer so that he is loyal, and Profit is also not affected..!!
This action U is denoted as U = (Atr, X—Y)
There are three modules of the system.[12]

A. Customer

In this module customer perform many operations as customer login, he can view various products, purchase products, add products to cart, Customer can recommend products to his/her friends. Here the system will generate static and dynamic recommendation according to the user's actions. The customer perform operations related to products. Customer is the basic factor of the recommendation system.[12]

B. Service Provider

The service provider is main module of the recommender system as it implements the decision tree. This module is responsible for generating probability of Loyal/Unloyal customer, which becomes input for the Admin module. The system will secured the user data from the service provider using Homomorphic encryption. This is fully based on the Support team Vs Customer module. This module has authority to view all customer details and can view all feedback information.

The service provider can prepare the customer profile for all customers. This Customer profile is containing loyal customer and unloyal customer. This module has sub modules as display customer details, display feedback information; build customer profile, profit calculation, listing action sets etc.[12]

C. Privacy Service Provider

Privacy service provider module distribute the customers as Loyal/Unloyal based on the probability generated from service provider module and perform action in order to retain the customers. Privacy service provider module determines the cost related to each action and the net profit associated with the action. Data wiil be visible to admin in decrypted format. This is the controller of the all modules. Admin has a full authority to view all customer information. Admin select the

correct suggestion from the customer, support team suggestion for increase customer in the management process and also apply the particular action for increase the accuracy.

This module has sub modules as view customer feedback, view support team suggestions, view action set, selecting action set, apply action, add new product ,view graphical reports etc.[12]

ACKNOWLEDGMENT

I express my gratitude towards project guide Prof.Amrit Priyadarshi and Prof. S. S. Bere, Head, Department of Computer Engineering, Dattakala Group of Institution, Faculty of Engineering, and Pune who guided and encouraged me in completing the project work in scheduled time. I would like to thanks our Principal Dr. S.S. Ragit for allowing us to pursue my project in this institute.

I thanks to Prof. Amrit Priyadarshi, ME Coordinator, for their guidance and for being a constant source of support. No words are sufficient to express my gratitude to my family for their unwavering encouragement. I also thank all friends for being a constant source of my support.

REFERENCES

[1] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," IEEE Trans. Knowl. Data Eng., vol. 17, no. 6, pp. 734–749, Jun. 2005.

[2] H. Polat and W. Du,"Privacy- Preserving collaborative filtering using randomized perturbation techniques," in proc. ICDM,2003,pp 625-628.

[3] H. Polat andW. Du, "SVD-based collaborative filtering with privacy," in Proc. 2005 ACM Symp. Applied Computing (SAC'05), New York, NY, 2005, pp. 791–795, ACM press.

[4] S. Zhang, J. Ford, and F. Makedon, "Deriving private information from randomly perturbed ratings," in Proc. Sixth SIAM Int. Conf. Data Mining, 2006, pp. 59–69.

[5] R. Shokri, P. Pedarsani, G. Theodorakopoulos, and J.-P Hubaux, "Preserving privacy in collaborative filtering through distributed aggregation of offline profiles," in Proc. Third ACM Conf. Recommender Systems (RecSys'09), New York, NY, 2009, pp. 157– 164, ACM.

[6] F.McSherry and I. Mironov, "Differentially private recommender systems: Building privacy into the net," in Proc. 15th ACM SIGKDD Int.Conf. Knowledge Discovery and Data Mining (KDD'09), New York, NY, 2009, pp. 627–636, ACM.

[7] R. Cissée and S. Albayrak, "An agent-based approach for Privacy preserving recommender systems," in Proc. 6th Int. Joint Conf. Autonomous Agents and Multi agent Systems (AAMAS'07), New York, NY, 2007, pp. 1– 8,ACM.

[8] Z. Erkin, M. Beye, T. Veugen, andR. L. Lagendijk, "Privacy enhanced recommender system," in Proc. Thirty-First Symp. Information Theory in the Benelux, Rotterdam, 2010, pp. 35–42.

[9] Z. Erkin, M. Beye, T. Veugen, and R. L. Lagendijk, "Efficiently computing private recommendations," in Proc. Int. Conf. Acoustic, Speech and Signal Processing (ICASSP), Prague, Czech Republic,May 2011, pp. 5864–5867, 2011.

[10] Casino, F. Domingo-Ferrer, J. Patsakis, C. Puig, D. Solanas, A., "Privacy Preserving Collaborative Filtering with k-Anonymity through Micro aggregation", e-Business Engineering (ICEBE), 2013 IEEE 10th International Conference on 11-13 Sept.

[11] I. Damgård, M. Geisler, and M. Krøigaard, "Efficient and secure comparison for on-line auctions," in Proc. Australasian Conf. Information Security and Privacy (ACSIP 2007), ser. LNCS, J. Pieprzyk, H. Ghodosi, and E. Dawson, Eds., Jul. 2–4, 2007, vol. 4586, pp. 416–430, Springer.

[12] Archana R. Kale, A. S. Hiwale, "Generating Customer Relationship Management Efficiently using Homomorphic encryption",IPGCON-2015.