I.J. Modern Education and Computer Science, **2019, 9, 1-13**
Published Online September 2019 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijmecs.2019.09.01

H2E: A Privacy Provisioning Framework for Collaborative Filtering Recommender System

Muhammad Usman Ashraf*
Department of Computer Science, GC Women University, Sialkot, Pakistan
*Email: m.usmanashraf@yahoo.com

Mubeen Naeem, Amara Javed, and Iqra Ilyas[1]
Department of Computer Science, GC Women University, Sialkot, Pakistan
Email: {mubeenaeem234, amarajaved706 }@gmail.com
[1]Email: iqra.ilyas@gcwus.edu.pk

Received: 21 July 2019; Accepted: 26 August 2019; Published: 08 September 2019

***Abstract*—**A Recommender System (RS) is the most significant technologies that handle the information overload problem of Retrieval Information by suggesting users with correct and related items. Today, abundant recommender systems have been developed for different fields and we put an effort on collaborative filtering (CF) recommender system. There are several problems in the recommender system such as Cold Start, Synonymy, Shilling Attacks, Privacy, Limited Content Analysis and Overspecialization, Grey Sheep, Sparsity, Scalability and Latency Problem. The current research explored the privacy in CF recommender system and defined the perspective privacy attributes (user's identity, password, address, and postcode/location) which are required to be addressed. Using the base models as Homomorphic and Hash Encryption scheme, we have proposed a hybrid model Homomorphic Hash Encryption (H2E) model that addressed the privacy issues according to defined objectives in the current study. Furthermore, in order to evaluate the privacy level, H2E was implementing in medicine recommender system and compared the consequences with existing state-of-the-art privacy protection mechanisms. It was observed that H2E outperform to other models with respect to determined privacy objectives. Leading to user's privacy, H2E can be considered a promising model for CF recommender systems.

***Index Terms*—**Recommender system, classification, collaborative filtering, privacy, Privacy techniques and Medicine recommendation.

I. Introduction

The extent of information in the world is growing far-off more rapidly than our capacity to process it. Thousands of new articles and blogs posted each day. An extensive collection of applications including recommendations in web search, books, movies, music, restaurants, food, apparels, vehicles, targeted advertisements, medicines, news, potential customers for companies and many more [2]. A Recommender System (RS) is one of the best technique that handles the information overload problem of Information Retrieval by proposing users with applicable and appropriate items. RS is intelligent enough to predict for a user his preference of one item over another. It is a property of RS that enable to give personalized recommendations to users. RS takes into version a grouping of multiple aspects to provide worthy recommendations. In other words, RS is either tools, techniques or applications that give specific suggestion about a user's interest. Help to decide on what to buy, what stocks to purchase etc. Apply in the variety of applications like a research paper, articles, movies, dramas, YouTube video and in other E-commerce websites.

These systems are mostly used by e-commerce websites to improve the user experience and thereby benefiting the stores. The system is able to convert browsers to buyers and cross-sell more items by means of suggestions while shopping. It increases user loyalty by enabling them to purchase items in fewer clicks and also providing frequent customers with good deals and offers. In short, an RS is able to attract the interest of the customers by providing them with fast and accurate recommendations. The first research paper in recommender systems came out in the mid-1990s [1] and since then research in this area got diversified and various approaches were introduced to present better recommendations.

There are three major classifications of RS that are collaborative filtering, content-based filtering, and hybrid filtering [3]. Content-based filtering, also sometimes known as cognitive filtering, recommend item according to the content of the user profile. If the user enters wrong information in their profile, then the RS not provide an accurate recommendation to an active user. More accurate user fills their profile, more accurate recommendation provided by a system [19]. On the other hand, Collaborative Filtering (CF) does not require any information about the users themselves. It recommends items based on users' past behavior. A system that

combines any two types of a recommender system to provide an accurate recommendation to an active user called hybrid filtering RS that takes the advantages of two recommender systems [29]. The detail classification is demonstrated by figure1.

A CF recommender system is a method of making a prediction about the interest of the user by collecting the preference of many other users. CF is a very powerful method for providing an accurate recommendation about user preference. CF technique can be distributed into two categories: memory-based CF and model-based CF. The implementation of a memory based system can either be item-based or user based [6]. In user-based CF, find other users whose previous rating is similar to that of the active user and use their rating on another item to predict what the current user will like [3]. It requires rating matrix and similarity function that calculate the similarity between two users. In item-based CF, recommend an item to the user that is similar to the user's highly preferred items. Cosine similarity matrix or conditional probability is used to compute item-item similarity. CF is a very powerful method for providing an accurate recommendation about user preference.

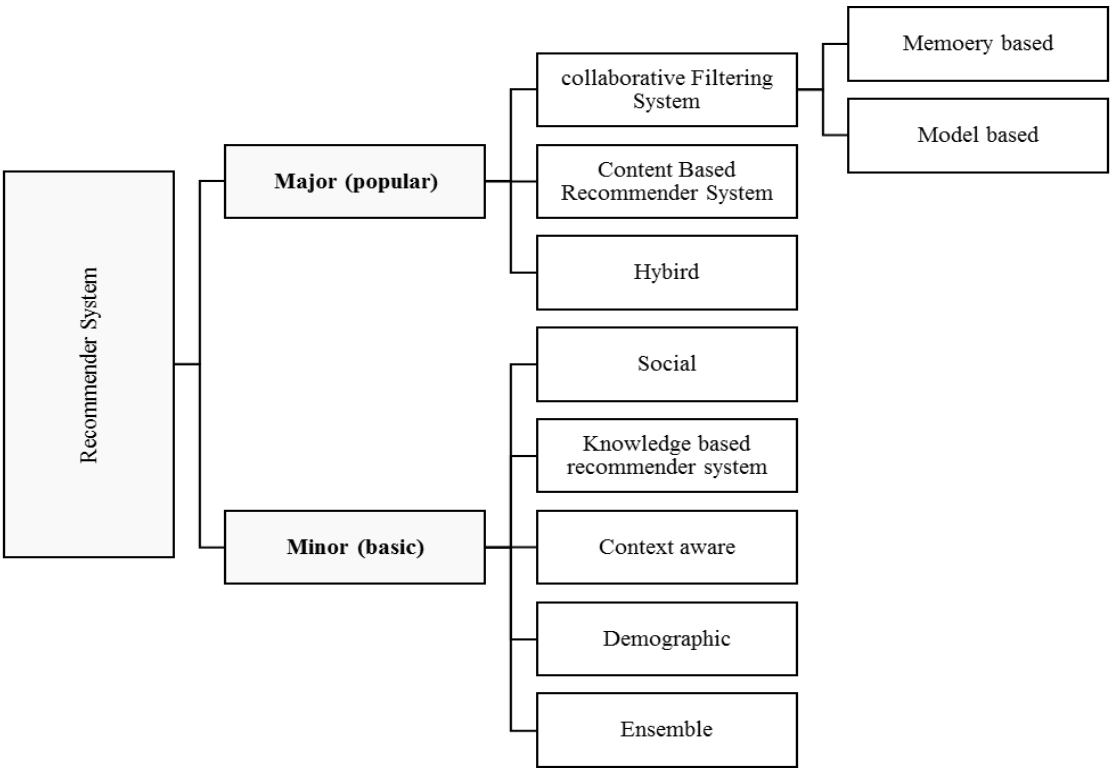

Fig.1. Classification of recommender systems

A. Collaborative filtering system

The collaborative approach [47] utilizes the proposal from different clients whose decisions are like the objective clients (for example client for whom the suggestion is made) Privacy. The clients with comparable decisions are named as a neighbour. Therefore, two noteworthy undertakings are being performed in community-oriented sifting; 1) finding the neighbour of a client and 2) investigating the inclinations of the neighbour of an objective client or client. The neighbour of a client can be framed by breaking down the past acquiring conduct of the client and computing the similitude scores between the decisions of these clients [48].

a. Model-based Collaborative filtering system:

The rating is used in the model-based recommendation. A model based [49] on the dataset of ratings. In other words, we extract some information from the dataset and use that as a "model" to make recommendations without having to use the complete dataset every time [50].

b. Memory-based collaborative filtering system:

Memory-based algorithms approach the collaborative filtering problem by using the entire database. As described by Breese et. Al [48], it tries to find users that are similar to the active user (i.e. the users we want to make predictions for) and uses their preferences to predict ratings for the active user [50]. Memory-based CF is divided into user-base [51] and item-based [52].

B. Content-based system

It's basically named an outgrowth and continuation of data separating research. In this framework, the articles are for the most part characterized by their related highlights. A substance based recommender [47] learns a profile of the new client's interest's dependent on the highlights present, in items the client has appraised. It's essentially a watchword explicit recommender framework

here catchphrases are utilized to portray the things. Along these lines, in a substance based recommender framework [48] the calculations utilized are with the end goal that it prescribes clients comparative things that the client has preferred before or is inspecting right now.

C. Hybrid Recommender System

Combining any of the two systems in a manner that suits a particular industry is known as Hybrid Recommender system [53, 54]. This is the most sort after Recommender framework that numerous organizations care for, as it joins the qualities of more than two Recommender framework [47] and furthermore kills any shortcoming which exists when only one recommender framework is utilized.

a. Weighted Hybrid Recommender

In this framework, the score of a suggested thing is figured from the aftereffects of the majority of the accessible proposal methods present in the framework. For instance, P-Tango framework consolidates communitarian and substance based suggestion frameworks giving them equivalent load in the beginning, however continuously changing the weighting as expectations about the client evaluations are affirmed or disconfirmed Paszzani's mix half breed doesn't utilize numeric scores yet rather treats the yield of each recommended as a lot of votes, which are then joined in an accord conspire.

b. Switching Hybrid Recommender

Switching Hybrid Recommender, switches between the suggestion systems dependent on specific standards. Assume in the event that we join the content and collaborative based recommender frameworks [47], at that point, the exchanging mixture recommender would first be able to send substance based recommender framework and on the off chance that it doesn't work, at that point, it will convey synergistic based recommender framework.

c. Mixed Hybrid Recommender

Where it's conceivable to make countless suggestions all the while, we ought to go for blended recommender frameworks [47]. Here suggestions from more than one procedure are introduced together, so it the client can browse a wide scope of proposals. The PTV framework, essentially a prescribed program to propose clients for TV seeing, created by Smyth and Cotter is utilized by a dominant part of the media and amusement organizations.

D. Demographic-based Recommender System.

This framework intends to order the clients dependent on traits and make suggestions dependent on statistic classes. Numerous ventures have adopted this sort of strategy as it isn't so mind-boggling and simple to actualize. In Demographic-based recommender framework, the calculations first need appropriate statistical surveying in the predefined district went with a short review to accumulate information for classification. Statistic systems structure "individuals to individuals" relationships like community-oriented ones, however, utilize various information. The advantage of a statistic approach is that it doesn't require a background marked by client evaluations like that in collective and substance based recommender frameworks.

E. Knowledge-based Recommender System

This type of recommender framework endeavors to propose articles dependent on surmising about a client's needs and inclinations. Information put together proposal [47] works with respect to utilitarian learning: they know about how a specific thing meets a specific client need, and can in this way reason about the connection between a need and a conceivable suggestion.

F. Social based recommender system

Social Recommender Systems (SRS) are recommender systems [47, 45] that target the social media domain. They aim at coping with the social overload challenge by presenting the most relevant and attractive data to the user, typically by applying personalization techniques.

G. Context-aware

Context-aware recommender systems (CARS) [47] generate more relevant recommendations by adapting them to the specific contextual situation of the user. This article explores how contextual information can be used to create more intelligent and useful recommender systems.

H. Ensemble

Consolidating any of the at least two than two frameworks in a way that suits a specific industry is known as group Recommender framework [47]. This is the most sort after Recommender framework that numerous organizations take care of, as it consolidates the qualities of more than two Recommender framework and furthermore takes out any shortcoming which exists when only one recommender framework is utilized.

While there are many problems in RS such as:

- When the user creates an online profile, for using social media or for E-Commerce application, they are usually not aware of the privacy of personal information.
- The user doesn't know about the level of information that a service provider collects, and how this information can be further use.
- Service provider or some internal employees can misuse or send this personal information to other service providers [5].
- Even when we rate some articles online, a third party can easily access our personal information data.

That's why we focus on privacy issues and how to provide privacy in CF recommender system. Our primary focus on user profile so that user profile not is easily accessible to the third party and not reveal the personal information.

In order to achieve the privacy goals, there are two main objectives of current research, first one is to find out the privacy factors that need privacy in CF recommender system so that user's personal information cannot be revealed and the second objective is to propose a new technique to minimize the privacy issue.

The structure of our paper is as follows: First, we present the classification of a recommender system. In Sect. 2 we describe some related work of recommender system. In Sect. 3, we identify different privacy attributes. In Sect. 4, we define approaches for provisioning privacy and describe briefly our proposed model H2E. In section 5 we present implementation process and results of our implementation. Then, we provide some discussion and recommendation about our research in Sect. 6. Finally, in the last sector, we close our research work with a short summary.

II. Related Work

Several state-of-the-art methods have been proposed that are utilized in a collaborative filtering recommender system framework for provisioning privacy to end user. These techniques include Elgamal Homomorphic Encryption [7], Data obfuscation, Cryptographic method, Privacy-Preserving Cryptographic Protocols with Server, Privacy Preserving Cryptographic Protocols without Server, Randomization-based and k-Anonymous approach [9]. There are also likewise non-technical approaches that are utilized for provisioning privacy to end users like awareness, law, and regulation. The detail of these techniques are given below:

A. Elgamal Homomorphic Encryption

User's private or sensitive information can be abused in light of inquisitive directors in online applications. Presently multi day's a large portion of individuals are utilizing on the web administrations for day by day exercises, which require offering individual data to the specialist co-op. A few models are informal organizations and web-based shopping. Dynamic client cooperation is a must. To beat this issue of dynamic cooperation of user's Erkin present homomorphic encryption plots and secure multiparty calculation (MPC) strategies for protection upgraded recommender framework by utilizing a semi confided in an outsider. Insights concerning the client's appraisals, closeness esteem estimation to a limit and created proposals are altogether kept avoided SP (service provider), PSP (private service provider) and different clients. To begin with, the clients conceal their own information by encryption and send it to the specialist organization. Also, to produce proposals the specialist co-op and the PSP run a cryptographic convention without connecting with the end users. For this goal, Paillier framework is used previously. This cryptosystem is utilized to encode the privacy-sensitive information of the end users [4].

The technique of collaborative filtering is utilized with the Paillier framework which incorporates the following:

- Compare the likenesses between a user and a different user.
- The comparative clients are chosen by contrasting their comparability with other. By normal rating of most comparable clients, the recommendations are produced.

In Paillier framework user's dynamic support is must, which makes the framework very tedious just as unpredictable. This is on the grounds that the user needs to perform all encryptions and decryptions a lot of times, which makes the framework expensive. Elgamal calculation is a lot easier and simple with an independent private key which makes it wasteful [5].

By large review demonstrates that the framework shows the Multiparty Computation system that gives a few issues or impediments. Accordingly, to maintain a strategic distance from these issues we can use two calculations that help us to profit for the individual and business and contrasted with existing private recommender systems techniques, this framework is increasingly secure, simple and efficient. This technique gives privacy to the user a specific level yet at the same time, privacy preserving techniques need to improve regarding precision/accuracy. Later on, we work remove away at the change in transmission and privacy that relies upon picked parameters.

B. Randomization-based

Randomization data is the way toward making something irregular. Information is veiled utilizing RPTs, in which numeric rating based collective separating plans. Information is aggravated utilized RPTs in which twofold evaluating based CF framework. [12] In RPTs have a random number they are appropriated with zero methods and a standard deviation. In information mask to get a genuine double appraising, at that point, the rating is as indicated by some of the gatherings. If the RN is huge than the limit, at that point the double evaluating is data switched. Random number circulations can be utilized to create an irregular commotion. All users can use a similar random number appropriation with the fixed estimations of protection control parameters.

First decide the number of gatherings and a limit. To mask the genuinely paired appraisals, the evaluations are assembled by the number of gatherings. Information in each gathering is then freely veiled. An irregular number is picked for each gathering and the arbitrary numbers are contrasted and the picked limit. In the event that the arbitrary number is bigger than the limit, parallel evaluations are turned around (1s are changed into 0s are changed into 1s) and sent to the CF frameworks [26]. For both of RRTs and RPTs, the number of arbitrarily chose unrated thing cells relies upon the measure of the unrated thing cells [27].

The limitations are lossy change, forbids characteristic grouping on information and protection relies upon picked parameters. Their future research course incorporates giving assessment structures to protection safeguarding synergistic separating plans. Likewise,

structures ought to be planned by thinking about security, privacy, performance, and robustness.

C. K-Anonymous

K-Anonymity is a property of an informational collection, generally utilized so as to depict the informational collection's dimension of obscurity. A dataset is k-unknown if each mix of personality uncovering qualities happens in at any rate k various lines of the informational collection. Comprises of a lot of k objects which implies more than one items from which a focused on an article is one which is hard for an aggressor to locate that one k object from a similar arrangement of k objects. Keep up factual qualities on information for little K values. It comprises of k questions so for an interloper it's hard to discover the k-1 object. It is anything but difficult to actualize [9].

The limitations are denied grouping on information, lossy change and privacy rely upon picked parameters. Future work will concentrate on two unique directions. The first is to improve the effectiveness of our technique so as to be actualized in a decentralized plan. The second directions are to dissect the impact of attribution method on privacy and suggestions quality, at that point contemplate systems of trust and proficient ascription arrangement.

D. Data Obfuscation method

Data Obfuscation method is likewise called information concealing in which information is mixed to square unapproved to get to sensitive information/material. Data obfuscation methods are utilized to forestall the interruption of private and sensitive online information [17]. Most well-known methodologies used in data obfuscation method is randomized annoyance based method. RPTs are anything but difficult to actualize and saves factual qualities to ensure private information. In which user preference can be spoken to utilizing numeric or binary rating. Utilizing random perturbation techniques for information mining in the region of privacy assurance. Prescribed framework quality assessment measurements, the normal supreme mistake is the most widely recognized estimation technique [11]. Randomized perturbation techniques are right off implemented by Polat and Du RPTs so as to accomplish confidentiality in Collaborative Filtering frameworks. The randomized perturbation methods are used for the user information gathering and can produce a recommendation with not too bad exactness. This methodology gives privacy to a user to a specific dimension yet at the same time, the privacy-preserving technique needs to upgrade as far as exactness. Later on, we take a shot at the lousy change and protection relies upon picked parameters.

E. Cryptographic method

Julius Caesar (100BC) use cryptography for the first time. Erkin et al. proposed harmonic encryption a technique utilized in the recommender framework (2012).kaleli proposed Concordance-based to use the arrangement of Harmonic encryption. Cryptographic is a strategy to scramble the information for secure correspondence. Its concern with Confidentiality, Integrity, and Authentication. It is utilized to encode private information for delivering a recommendation. The concordance-based arrangement uses the harmonic encryption to give the predictions without taking a chance with privacy [16].

Users rating are scrambled the open key of confided in experts and submitted to an open server. User can scramble them all ratings and the server registers the normal and likeness among things utilizing symphonious properties and enables all the user to decrypt which save their secret information. It passes the messages on item comparability among user and server which make it secure [14].

In the cryptographic technique, there is a huge amount of information that is required to encrypt and decrypt which require an enormous computational expense and furthermore it makes the framework more complex. In spite of the fact that this methodology gives privacy to the user a specific dimension yet at the same time, privacy preserving techniques need to improve regarding exactness and dependability. The privacy-preserving techniques are assembled as brought together v/s decentralized as future work.

F. Privacy-Preserving Cryptographic Protocols with Server

Privacy-preserving cryptographic protocols with a focal server expect to utilize the centralization offered by the service provider while utilizing secure two-party computation and encryption to guarantee the privacy of the users. The centralized structure is preserved yet neither the operator nor the organization can connect the user's rating to the items. User rating is encoded the public key of confided and submitted to a server. A central server goes about as a middle person between the users and is accountable for consolidating the outcomes given by various users. While wanting a recommendation, a user sends an encoded solicitation to the central server. The server disperses this solicitation to different users that can take a shot at the solicitation by utilizing the homomorphic properties of the cryptosystem. A protected two-party computation at that point decides for every user if their data ought to be incorporated into them or not. The central server at that point consolidates the (still encoded) results to produce the recommendation [14]. The disadvantage of these techniques (that include a layer of encryption) is effectiveness. The homomorphic tasks and secure two-party calculations are in every case costlier than their unprotected partners. Actually, the error is frequently gigantic. These outcomes in poor proficiency and adaptability for these conventions. In spite of the fact that this strategy gives protection to user’s specific dimension yet at the same time, the privacy-preserving technique needs to upgrade as far as precision, reliability, effectiveness as future work.

G. Privacy-Preserving Cryptographic Protocols without Server

Privacy-preserving cryptographic without a central server plan to evacuate the trust that is put in specialist organizations by expelling them from the image. The users enter the value for the collaborative filtering process which at that point uses to register the model for further preparing and give proposal as indicated by the user's preference [14]. It includes plenty of users which may cause deferral and low exactness. Utilize that strategy where create suggestion with the presence of privacy and dynamic user's participation isn't necessary. There are numerous strategies accessible where dynamic user's participation isn't necessary like homomorphic encryption method [15].

The disadvantage is the association of numerous users that are required to make (the model for) the recommendation. These users need to interact with one another, yet not all users will be accessible in the meantime. This can prompt significant postponements or lost precision. Despite the fact that this technique gives privacy yet it needs to think about the accuracy as a future work.

After analyzing the described state-of-the-art methods for privacy provisioning in CF recommender systems, we analyzed that a new approach is required that consider all the privacy factors discussed in the following section. In our proposed model H2E, user's profile is fully secure from service provider and their internal users so that user data not be misused for any purpose. This is the main contribution of our research.

III. Privacy factors for CF Recommender System

In this section we have described the primary privacy attributes that have an important role in user privacy. Many users feel hesitation to use e-commerce websites. A mobile application user interacting with CF recommender system can be safe if the mentioned factors are secured. The core privacy factors of CF recommender systems has been discussed in table 1 as follows.

Table 1. Privacy attributes for CF recommender systems

Privacy factors	**Description**
Postcode, location, address	"Every single zone in the nation has been distinguishing by the postcode". Postcode is spoken to by 4 letters. "Each area can be followed by knowing these 4 letters".
Name	F.Name and L.Name. "Characters speak to the name"
Type of device	"Each device has there on IP address". The device can be followed by its IP address.
Gender	"Male or Female". It speaks to characters.
Age	"Continuously indicated by the birth". It is spoken to by the moth and year. Age can be followed by birth.
Status	"Similarly as with the statistic variable of age, we were unfit to locate any quantitative survey of the relations between identity factors and ethnic status."
Social Security Number	An SSN is a nine-digit number that is by and large issued to U.S. natives, legal perpetual occupants, and certain (working) non-immigrants.

IV. Proposed Homomorphic Hash Encryption (H2E) Model

We propose a new model of a privacy-preserving collaborative filtering recommender system, by using the base models homomorphic technique and Hash function encryption technique proposed a Hybrid model homomorphic hash encryption (H2E), which allows the computations required for recommendations in a dispersed manner and preserves user privacy without compromising recommendation accuracy and efficiency. By using the hybrid model we achieve privacy to all the factors that are the main concern of users. We introduce the privacy protocol by hash function encryption which is based on key generation [1]. We assume a semi-trusted server named ‘‘recommender server" whose task is to perform the computations for the recommendation on encrypted data. We propose different privacy protocols for item average and similarity computations as well as recommendations generations by which the privacy of users is preserved. Specifically, our main contributions are:

- An efficient privacy-preserving item-based recommender system to protect user privacy during the recommendation process.
- Privacy-preserving item average and similarity computation protocols to calculate averages and similarities among the items without compromising user ratings. Moreover, which items have been rated are also hidden during these processes.

H2E Algorithm

Input: PNum ← phone number
P ← password
S ← symptoms
Output: R ← Recommendation
Declaration: T ← timer, D ← diseases, M ← Medicine.
Key Generation:
Key generation involves one to four steps.
(1) User enters personal information.
(2) Check validation.
(3) Code send to user mobile and add T.
(4) Login.
(5) Password Hash.
Recommendation Computing:
Recommendation generation involve five to nine steps.
(6) Get S using Get () method.
(7) Find similarities between S and D.
(8) Count Sid Count>= 3 Did return.
(9) SELECT *FROM medicine WHERE Did LIKE Sid.
(10) Recommend M to user.
(11) User rates that M.

Our proposed H2E model work as user sign up into system then validation will be performed. After that firebase verification is performed. Then password and phone no are entered to log in. Password has due to hash function encryption and the hash of the password is

stored in the database. First, a user selects the symptoms from the given list of symptoms then the system finds out the similarities between symptoms and disease by using their identities. Secondly, calculate similarities between the given diseases and medicine. Then recommended top rated items or medicine to the active user on the bases of their interest or preferences. User can rate the medicine and also see top-rated medicine at the same time without affecting the performance of the recommender system. The flowchart of the proposed model has been presented in figure 2 as follows.

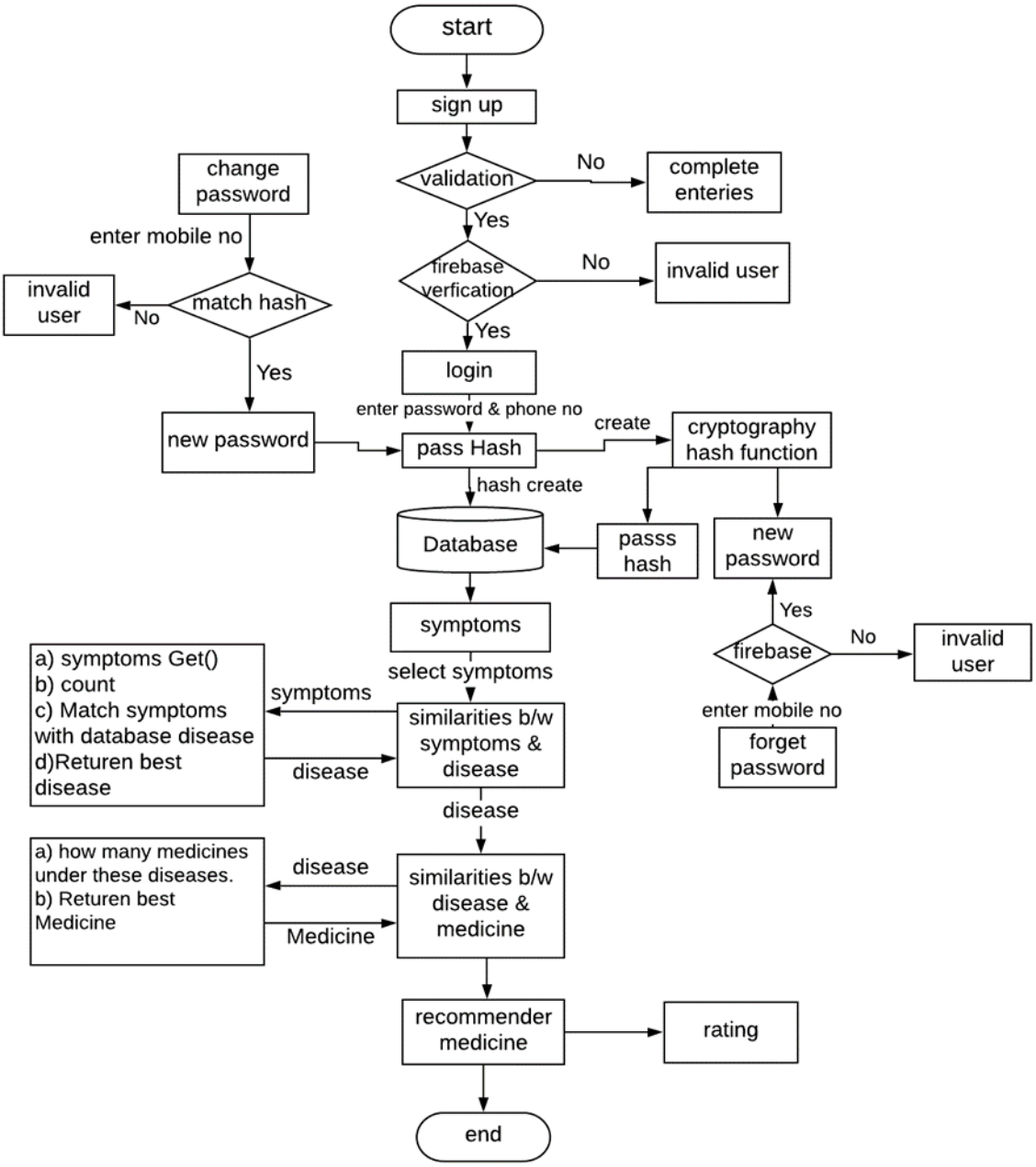

Fig.2. Workflow of H2E Model

V. Implementation and Results

The specified application is implemented by using different tools. For this purpose, the system description may include Processor of Intel(R) Core ™i7-8550U CPU @ 1.80GHz 1.99GHZ, installed memory (RAM), 16.0GB and 64-bit operating system, the x64-based processor is used.

The application based on Flask framework. It uses JavaScript. MySQL is used as a Database Management System (DBMS) for the application. JavaScript is used for validation. Similarly, MS Excel is used for data pre-processing and draw.io is used as a case tool. The

application's implementation work is divide by using Critical path method (CPM).

Our CPM analysis carried out as follows in table 2.

Table 2. Critical path method

Activity	Time (weeks)	Predecessor
Data collection (A)	2	-
Database design (B)	1	-
Data preprocessing(C)	1	A,B
Registration of user(D)	2	C
Recommendation implementation(E)	2	D
Privacy implementation(F)	2	E
Front end design(G)	3	E,F
Testing(H)	1	G
Documentation(I)	3	D

It is noted that the application was completed in 17 weeks, which is within a semester. So, the project proved feasible in terms of schedule. All the activities are crucial for the implementation of the application.

To carry out implementation first requirements gathered, as they are a critical part of any application. There are two types of requirements, functional and non-function, given below in Table 3.

Table 3. Functional and non-functional requirements

Sr. NO.	Functional requirements	Non-functional requirements
1.	User registration	(1) User can only register with a valid mobile number. (2) Only one account can be created with a mobile number. (3) The user name must contain a minimum of four characters. (4) Password length must be eight letters with one digit must. (5) The user should fill all requirements for registration otherwise, the system not entertain that particular user.
2.	Login	(1) User login with a valid phone number and password as user enter at the time of registration.
3.	Select symptoms	(1) User must select valid symptoms. (2) The user selects at least three symptoms for recommendation according to other similar users.
4.	Rate medicine	(1) User can only rate medicine on the scale of 1 to 5, one rating represents worst and five ratings represent best medicine and quality. (2) User can only rate medicine just once.

Since both individual collaborative filtering has their own limitations which can be minimized in an application if both of the algorithms are used.

Listing 1. Recommendation generation

```
$sql="select * from `medicine` WHERE `did` LIKE '%". $id ."%' ";

$result=mysqli_query($this->dbConnect(), $sql);
if(mysqli_num_rows($result) > 0){
while ($res = mysqli_fetch_assoc($result)) {
$arr[] = array("id"=>$res['id'], "title"=>$res['name'],
"side_effect"=>$res['side_effect'],
"indication"=>$res['indication'],
"alternative"=>$res['alternative'],
"contradiction"=>$res['contradiction']);
}
$this->response($this->json($arr) , 200);
}else{
$arr[] = array("status"=>"False" , "message"=>"Try again with real symptoms");
$this->response($this->json($arr) , 200);  $sql="select * from `medicine` WHERE
`did` LIKE '%". $id ."%' ";
```

Listing 2. User validation

```
{
Pattern pattern;
Matcher matcher;
final String PASSWORD_PATTERN = "^(?=.*[a-zA-Z])(?=.*\\d)(?=.*[!@#$%^&*()_+])[A-Za-z\\d][A-Za-z\\d!@#$%^&*()_+]{7,19}$";
pattern = Pattern.compile(PASSWORD_PATTERN);
matcher = pattern.matcher(password);
return matcher.matches();
}
```

The results carried out by doing a feasibility analysis. We analyse schedule feasibility by using Critical path method. CPM was used for identification of critical tasks an also calculated the relationship between tasks. The application uses XML to display content in the mobile, style resource file is used for applying styles and .java file is used for making an interactive application. At the server side, it uses PHP to implement the logic. It requires a server, client and an internet connection to work properly. It sports all devices except the iPhone. All of the technologies required by the application are available and easily accessed, therefore it was strong-minded technical feasibility. Our mobile application used 2-tier architecture where the clients of application are the end user who gets recommendations and rate that medicine. The server maintains the records of all user, medicines, diseases, symptoms, user's rating, and user history. Our application access anywhere with the help of internet connection and it is easy to operate. That is why it is determined that our application is operationally feasible.

We check our system validation by using precision and recall model. The test date is used to test the System and the application obtained 76.5% precision and 61.7% recall.

Precision =positive rating/actual results Recall true (1)

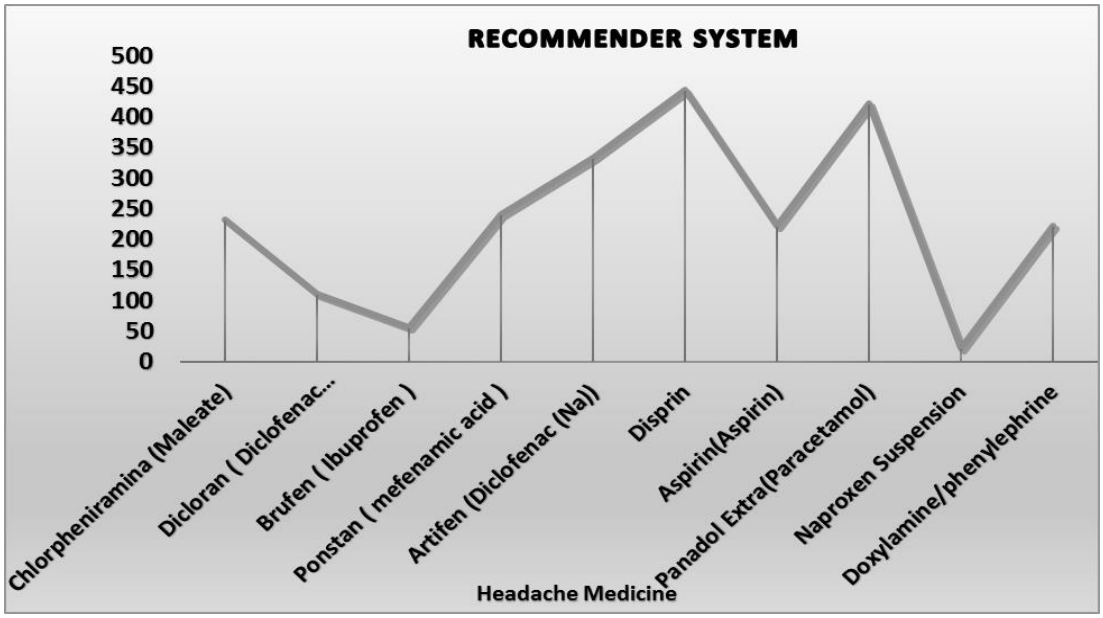

Fig.3. Recommendation according to users

The diagram illustrates the recommendation process. It shows that medicine with high rating recommends to all users. The rating of headache medicine Disprin is high as compare to others so it is recommended to use with their rating and of course user must select hat medicine that has been rated or like by many other users.

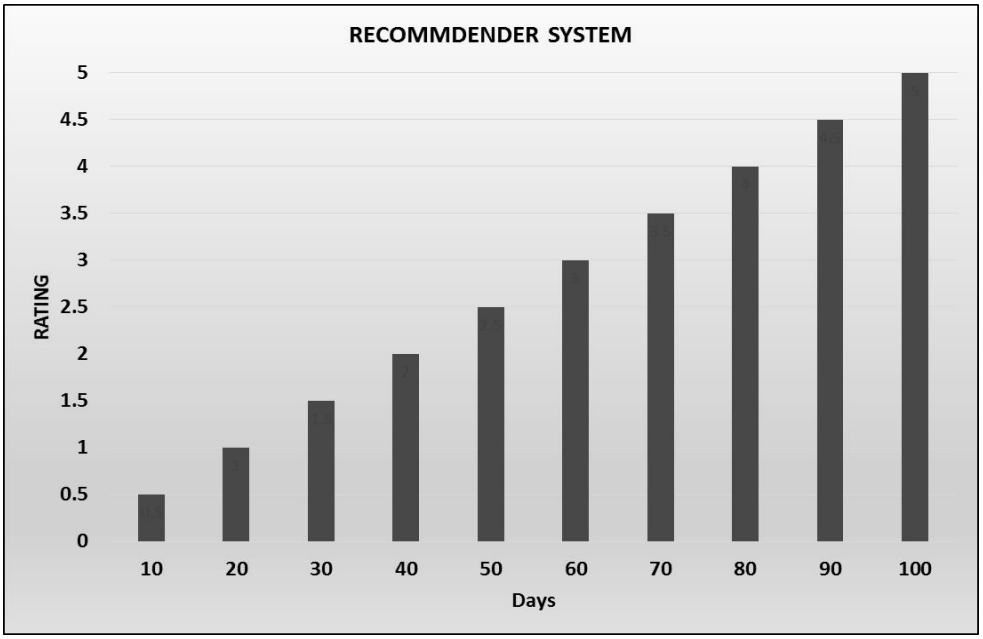

Fig.4. Rating of products

The diagram shows that initially, the quantity of user rating is very low. It is really difficult to recommend user because similar finding data is too short but with the passage of time the level of rating increase and system recommend user immediately in less and less time. In other words, the recommendation process grows faster as compared to the initial level.

In other words, the recommendation process grows faster as compared to the initial level. Comparison with previous research is shown in 7.

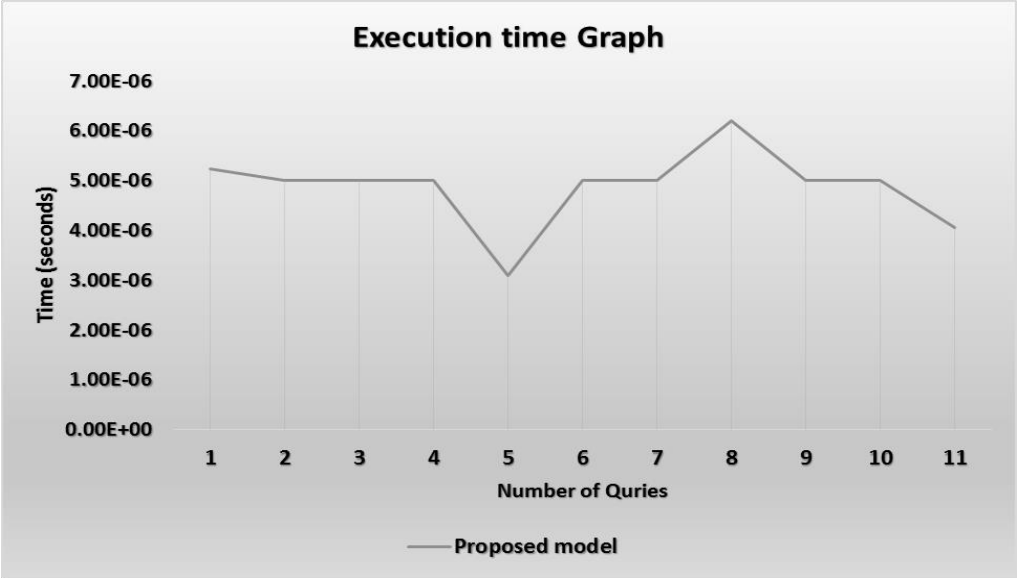

Fig.5. Execution time of H2E

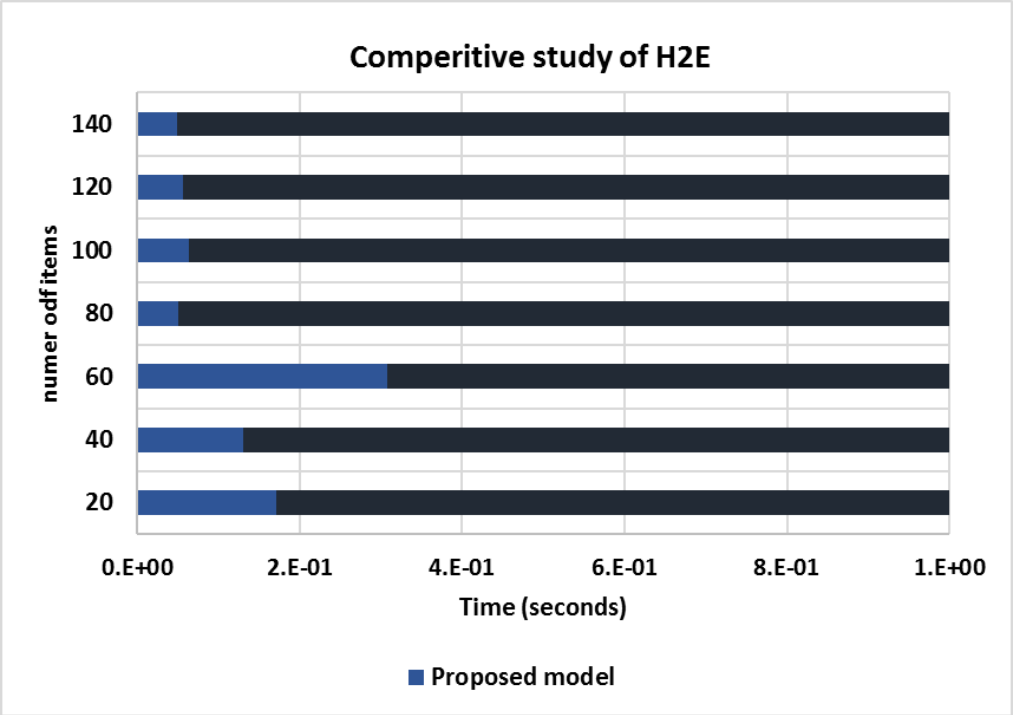

Fig.6. Comparative analysis of H2E

VI. Discussion and Recommendation

The primary objective of the current research was to protect user's privacy. Above mentioned state of the art approaches protect some user's privacy attributes for the provision of privacy [22]. Here, we elaborate all approaches according to some specific attributes given by them. Moreover, to protect user's rating from any unauthorized person due to privacy issues, we provide the best technique through which service provider and their internal employee do not access user profiles because ewe the password and without password nobody can access the profile of a user. The privacy requirements change from person to person. In the last few decades the work has been done on privacy but privacy issues still exist and user needs to secure their personal information in every situation. Sometimes, the data is not deleted from the server after ordering in e-commerce websites. That's why some privacy issues arise and users feel hesitation to use e-commerce websites. Data from the main server can be misplaced. The user needs privacy and we present the best privacy provisioning techniques which help a lot in achieving user's privacy requirements. No doubt, all these techniques helped a lot to get a good deal for preserving active user privacy but still, there are privacy issues. All above-mentioned techniques do not provide complete user satisfaction. We apply one of the

best technique for achieving our privacy goal by protecting the above-mentioned privacy factors. As we mentioned above that we use H2E for securing user data [23].

The experimental part focused on improving the privacy preservation by applying Homomorphic encryption technique and by using cryptography hash algorithm. Initially, the error of generated prediction is not good and accurate because in start user's rating rate is low and with the passage of time user's rating increase, hence it is easy to find the similar user and predict the preference of the particular user. Other results showed that extreme ratings had a stronger effect on the accuracy of the predictions than low or moderate ratings. This allowed us to conclude that the extreme ratings are important for the accuracy of CF recommendations, as they allow identifying the real preferences of the users.

These results introduce an important CF trade-off. From one hand, the results showed that the extreme ratings are important for the generation of accurate CF predictions. From the other hand, an experiment showed that the users consider extreme ratings in their profiles as more sensitive and prefer not to expose them by applying cryptography hash function which encrypts the password of the user into a hash which cannot be decrypted at all. Combination of these two conclusions highlights the trade-off between accuracy and privacy in CF indicates that there is no simple way to optimize both the accuracy of the recommendations and privacy of the users [28].

VII. Conclusion

Recommender System (RS) has been used in the last few years. Advanced techniques have been implemented to get a first-rate and modified RS. Yet, designers face privacy disputes due to which the user's personal information revealed out through a service provider who shares the user's information to another service provider for knowing their market position. In order to provisioning privacy for defined factors in CF recommender system, we have introduced a new Homomorphic Hash Encryption (H2E) model that enhanced the privacy when a user interacts with CF recommender system. In order to validate the privacy success rate, we implemented in real time medical application by using existing datasets. Moreover, the results were compared with existing state-of-the-art methods and noticed that the proposed H2E model outperformed by providing maximum privacy to users.

By future perspectives, the proposed H2E model should be investigated by implementing at large scale systems to know the level of privacy when a huge number of recommendations are required through frequent interactions.

References

[1] Polat, H. and Du, W., 2005. Privacy-preserving collaborative filtering. International journal of electronic commerce, 9(4), pp.9-35.

[2] Berkovsky, S., Eytani, Y., Kuflik, T. and Ricci, F., 2007, October. Enhancing privacy and preserving accuracy of a distributed collaborative filtering. In *Proceedings of the 2007 ACM conference on Recommender systems* (pp. 9-16). ACM.

[3] Park, D.H., Kim, H.K., Choi, I.Y. and Kim, J.K., 2012. A literature review and classification of recommender systems research. *Expert systems with applications*, *39*(11), pp.10059-10072.

[4] Kaleli, C. and Polat, H., 2010. P2P collaborative filtering with privacy. *Turkish Journal of Electrical Engineering & Computer Sciences*, *18*(1), pp.101-116.

[5] Yargic, A. and Bilge, A., 2017, July. Privacy Risks for Multi-Criteria Collaborative Filtering Systems. In *2017 26th International Conference on Computer Communication and Networks (ICCCN)* (pp. 1-6). IEEE.

[6] Sarwar, B.M., Karypis, G., Konstan, J.A. and Riedl, J., 2001. Item-based collaborative filtering recommendation algorithms. Www, 1, pp.285-295.

[7] Patil Maulik, Y. and Yeola, M., Generating Private Recommendations Using ElGamal Homomorphic Encryption.

[8] Zhan, J., Wang, I.C., Hsieh, C.L., Hsu, T.S., Liau, C.J. and Wang, D.W., 2008, August. Towards efficient privacy-preserving collaborative recommender systems. In *2008 IEEE International Conference on Granular Computing* (pp. 778-783). IEEE.

[9] Casino, F., Domingo-Ferrer, J., Patsakis, C., Puig, D. and Solanas, A., 2015. A k-anonymous approach to privacy preserving collaborative filtering. *Journal of Computer and System Sciences*, *81*(6), pp.1000-1011.

[10] Taziki, M., Differential Privacy in Recommenders. Parameswaran, R. and Blough, D., 2005. A robust data obfuscation approach for privacy preservation of clustered data. In *Workshop on privacy and security aspects of data mining* (pp. 18-25).

[11] Parameswaran, R. and Blough, D., 2005. A robust data obfuscation approach for privacy preservation of clustered data. In Workshop on privacy and security aspects of data mining (pp. 18-25).

[12] Batmaz, Z. and Polat, H., 2016. Randomization-based Privacy-preserving Frameworks for Collaborative Filtering. *Procedia Computer Science*, *96*, pp.33-42.

[13] Batmaz, Z. and Kaleli, C., 2017, October. Methods of privacy preserving in collaborative filtering. In *2017 International Conference on Computer Science and Engineering (UBMK)*(pp. 261-266). IEEE.

[14] Ekstrand, M.D., Riedl, J.T. and Konstan, J.A., 2011. Collaborative filtering recommender systems. *Foundations and Trends® in Human–Computer Interaction*, *4*(2), pp.81-173.

[15] Su, X. and Khoshgoftaar, T.M., 2009. A survey of collaborative filtering techniques. *Advances in artificial intelligence*, *2009*.

[16] Jeckmans, A.J., Beye, M., Erkin, Z., Hartel, P., Lagendijk, R.L. and Tang, Q., 2013. Privacy in recommender systems. In *Social media retrieval* (pp. 263-281). Springer, London.

[17] Bokde, D., Girase, S. and Mukhopadhyay, D., 2015. Matrix factorization model in collaborative filtering algorithms: A survey. Procedia Computer Science, 49, pp.136-146.

[18] Parameswaran, R. and Blough, D.M., 2007, November. Privacy preserving collaborative filtering using data obfuscation. In *2007 IEEE International Conference on Granular Computing (GRC 2007)* (pp. 380-380). IEEE.

[19] Akhil, P.V. and Joseph, S., 2017. A SURVEY OF RECOMMENDER SYSTEM TYPES AND ITS CLASSIFICATION. *International Journal of Advanced Research in Computer Science*, *8*(9).
[20] Patil Maulik, Y. and Yeola, M., Generating Private Recommendations Using ElGamal Homomorphic Encryption.
[21] Stekh, Y., Lobur, M., Artsibasov, V. and Chystyak, V., 2015, February. Methods and tools for building recommender systems. In *The Experience of Designing and Application of CAD Systems in Microelectronics* (pp. 300-305). IEEE.
[22] Friedman, A., Knijnenburg, B.P., Vanhecke, K., Martens, L. and Berkovsky, S., 2015. Privacy aspects of recommender systems. In *Recommender Systems Handbook* (pp. 649-688). Springer, Boston, MA.
[23] Jeckmans, A.J., Beye, M., Erkin, Z., Hartel, P., Lagendijk, R.L. and Tang, Q., 2013. Privacy in recommender systems. In *Social media retrieval* (pp. 263-281). Springer, London.
[24] Berkovsky, S., Eytani, Y., Kuflik, T. and Ricci, F., 2005, July. Privacy-enhanced collaborative filtering. In *Proc. User Modeling Workshop on Privacy-Enhanced Personalization* (p. 46).
[25] Gong, S., 2011. Privacy-preserving collaborative filtering based on randomized perturbation techniques and secure multiparty computation. *International Journal of Advancements in Computing Technology*, *3*(4), pp.89-99.
[26] Polatidis, N., Georgiadis, C.K., Pimenidis, E. and Mouratidis, H., 2017. Privacy-preserving collaborative recommendations based on random perturbations. *Expert Systems with Applications*, *71*, pp.18-25.
[27] Khusro, S., Ali, Z. and Ullah, I., 2016. Recommender systems: issues, challenges, and research opportunities. In *Information Science and Applications (ICISA) 2016* (pp. 1179-1189). Springer, Singapore.
[28] Ricci, F., 2014. Recommender systems: Models and techniques. *Encyclopedia of Social Network Analysis and Mining*, pp.1511-1522.
[29] Sielis, G.A., Tzanavari, A. and Papadopoulos, G.A., 2015. Recommender systems review of types, techniques, and applications. In *Encyclopedia of Information Science and Technology, Third Edition* (pp. 7260-7270). IGI Global.
[30] Wang, J., De Vries, A.P. and Reinders, M.J., 2006, August. Unifying user-based and item-based collaborative filtering approaches by similarity fusion. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 501-508). ACM.
[31] Neupane, P., 2016. *Restauraurant Recommendation System Based On Collborative Filtering* (Doctoral dissertation, Tribhuvan University).
[32] Deven Bansod, D. U. (2003 - 2019). Bringing MySQL to the web. Retrieved april 1, 2019, from phpmyadmin.net: https://www.phpmyadmin.net/
[33] developer, G. (n.d.). Android Studio. Retrieved March 1, 2019, from developer.android.com: https://developer.android.com/studio/
[34] Ho, D. (2019). Notepad++. Retrieved April 5, 2019, from notepad-plus-plus.org: https://notepad-plus-plus.org/
[35] Friends, A. (n.d.). XAMPP Apache + MariaDB + PHP + Perl. Retrieved april 5, 2019, from apachefriends.org: https://www.apachefriends.org/index.html
[36] "Genymotion" emulator for Windows, (version): 2.12.2 avaliable [online]: Retrieved from Genymotion 2.12.2 Android Emulator for PC Windows
[37] Goncharov, S.V., 2019. Using fuzzy bits and neural networks to partially invert few rounds of some cryptographic hash functions. *arXiv preprint arXiv:1901.02438*.
[38] Gheorghiu, V. and Mosca, M., 2019. Quantum cryptanalysis of symmetric, public-key and hash-based cryptographic schemes. *arXiv preprint arXiv:1902.02332*.
[39] Mašovic, S.H., Saracevic, M.H., Stanimirovic, P.S. and Krtolica, P.V., 2019. Computing Triangulations Of The Convex Polygon In Php/Mysql Environment.
[40] Jiang, J.Y., Li, C.T. and Lin, S.D., 2019. Towards a more reliable privacy-preserving recommender system. *Information Sciences*, *482*, pp.248-265.
[41] Meng, S., Qi, L., Li, Q., Lin, W., Xu, X. and Wan, S., 2019. Privacy-preserving and sparsity-aware location-based prediction method for collaborative recommender systems. *Future Generation Computer Systems*, *96*, pp.324-335.
[42] Kouki, P., Schaffer, J., Pujara, J., O'Donovan, J. and Getoor, L., 2019, March. Personalized explanations for hybrid recommender systems. In *Proceedings of the 24th International Conference on Intelligent User Interfaces* (pp. 379-390). ACM.
[43] Bengio, S., Dembczynski, K., Joachims, T., Kloft, M. and Varma, M., 2019. Extreme Classification (Dagstuhl Seminar 18291). Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.
[44] Huang, Y., Liu, H., Li, W., Wang, Z., Hu, X. and Wang, W., 2019. Lifestyles in Amazon: Evidence from online reviews enhanced recommender system. *International Journal of Market Research*, p.1470785319844146.
[45] Hassan, T., 2019, May. Trust and Trustworthiness in Social Recommender Systems. In *Companion* Proceedings of The 2019 World Wide Web Conference (pp. 529-532). ACM.
[46] Badsha, S., Yi, X. and Khalil, I., 2016. A practical privacy-preserving recommender system. Data Science and Engineering, 1(3), pp.161-177.

Authors' Profiles

M Usman Ashraf was born in Sialkot, Pakistan in 1988. He received the B.Sc. degree in Mathematics from The University of Punjab, Pakistan in 2007, M.S. degrees in Computer Science from the University of Lahore, Pakistan, in 2014 and currently doing Ph.D. degree in Computer Science from King Abdulaziz University, Jeddah, Saudi Arabia. He is Assistant Professor in the Department of Computer Science, GC Women University, Sialkot, Pakistan. He is also a member of Software Engineering group at King Abdulaziz University Jeddah, Saudi Arabia. From 2010 to 2014, he was a Senior Software Engineer (SSE) at Coeus Software solutions, GmbH. His research interests include High Performance Computing (HPC), Parallel Computing, Exascale Computing, Ubiquitous computing, Software Engineering, Location Based Service Systems and Recommender Systems.

Mubeen Naeem was born in Sialkot, Pakistan in 1998. She received the degree of Bachelors in information technology from GC Women University, Sialkot, Pakistan in 2015 and currently doing MS in computer science. She has also attended the conference of Cyber Secure Pakistan 2019 in March 12, 2019. Her research interest including recommender system and network security.

Amara Javed was born in Sialkot, Pakistan in 1997. She received the degree of Bachelors in information technology from GC Women University, Sialkot, Pakistan in 2015 and currently doing MS in computer science. Her research interest including recommender system and network security.

Iqra Ilyas was born in Sialkot, Pakistan. She received her bachelor's degree from the University of Punjab, in Lahore, Pakistan, in 2010, her M.Sc. (IT) from the University of Gujrat, in Gujrat, Pakistan in 2012, and her M.S (IT) from the University of Lahore, in Lahore, Pakistan, in 2016. She was a lecturer at GC Women University Sialkot Pakistan from 2013- 2017. Currently, she is serving as a network administrator at GC Women University Sialkot. Her research interests include Software Engineering, Data Mining, Cloud Computing and Artificial Intelligence, and Internet of Things.

How to cite this paper: Muhammad Usman Ashraf, Mubeen Naeem, Amara Javed, Iqra Ilyas, " H2E: A Privacy Provisioning Framework for Collaborative Filtering Recommender System", International Journal of Modern Education and Computer Science(IJMECS), Vol.11, No.9, pp. 1-13, 2019.DOI: 10.5815/ijmecs.2019.09.01

