FRIEND RECOMMENDATION SYSTEM BASED ON NON SPAM RATINGS

G.Abhinav Sai

Department of Electronics and Communication Engineering
Sreenidhi Institute of Science and Technology, Ghatkear, Telangana, India

CN Sujatha

Department of Electronics and Communication Engineering, Professor
Sreenidhi Institute of Science and Technology, Ghatkear, Telangana, India

Ch Saketh Reddy

Department of Electronics and Communication Engineering
Sreenidhi Institute of Science and Technology, Ghatkesar, Telangana, India

J Ranjith Kumar Reddy

Department of Electronics and Communication Engineering
Sreenidhi Institute of Science and Technology, Ghatkesar, Telangana, India

ABSTRACT: *In ever evolving social network system Friend recommendation system (FRS) is considered as a vital element. In the past few years, as the usage and familiarity of the social network sites is increasing, the number of FRSs is also increasing. Getting into details, the network systems recommend the people who has common feature like in the case of their friends. This commonality-based FRS is exact in cases where a physical or social feature like age, race, job, lifestyle or any other common factor exist. Personality types are an exemption in this. In any case, if a personality type is given, it is not necessarily needed that compatibility can develop with people of the same personality type. Chances are scanty. Through this project, we would like to present and gauge an FRS based on few personality traits and attributes. Initially some personality traits model is taken and hybrid filtering process is recommended. Later the users' harmony rating is done as a yard stick to measure and access the accuracy. This is how a personality-based social network site by using FRS functions and it is named as PersoNet. Results and users' ratings depict that the PersoNet performs collaborative filtering (CF)-based FRS. Moreover, there are lot of spam reviews that are available in social system. To make it more meaningful, we would be filtering the spam reviews and then make SPAN REVIEW SYSTEM based on 3 detection models: i) Average Content Similarity and ii) Rate Deviation and iii) Burstiness*
Keywords: Collaborative Filtering, PersoNet, Average Content Similarity, Rate Deviation, Burstiness.

I. INTRODUCTION

In today's world, social network applications are the main platform for building social circles, more than 3 billion active online social accounts are present in the online platforms. [1] Witnessing the current pandemic, people are interested to maintain a social circle in online social networking application only. Building a social circle on online platforms is a complicated task. Generally, for building the social circle, online users need to search the details of the user in applications by giving inputs such as Full name, Email, Contact number, etc. Manual searching of friends in online social applications cannot solve the problem of building a social circle. For building the social circle of the users, applications should concentrate on friend recommendations. In recommendation of friends in online social networking apps, we have observed few procedures which satisfy only the minimum requirements of the users. In current social networking apps using few procedures which are listed in the following.

1. Based on mutual friends or common friends.
2. Based on the Common Groups.
3. Based on the Contacts match.

The main intension of the Proposed work as we discussed many surveys proposed and motivated a friend recommendation system with various aspects. Our aim is to propose a system that has the caliber to distinguish friends considering personality of the users as a factor when segregating. Like previous concepts taken mutual friends, and profile attribute for input to calculating the friend recommendation same way we need to identify the few solutions for friend recommendation system. Those are,

→ Which data we consider for calculating the personality similarity of the users?

→ Which methodology we use for calculating the similarity of the users?

→ Need to build a system for identifying the similarity of the users without encouraging the spam data.

Objectives of the Proposed Work

To achieve our aims, the system that we came up with- Friend Recommendation System, which recommends the friends based on the calculation of the similar personality. Here we are considering the ratings of the users to calculating the personality similarity of the users. We can take any domain of the opinion data; in this project we are considering the E-commerce product related opinions are we are considering. Based on the harmony ratings of the users we define similarities of the users. To calculating the similar behaviors of the users we are implementing the Collaborative Filtering algorithm for the detection of harmony ratings. In our system Friend Recommendation System, we are considering the opinions main aspect to friend recommendation. In other hand it is very important to calculate the spam reviews and ratings, so that we also survey on the spam detection of the opinions. In this system we deployed three spam detection models to identify the spam review in terms of the review text, Meta data of the review like date of the review posting and rating behavior. Three spam models we listed in the following.

1. Content Similarity
2. Early Time Frame
3. Rate Deviation

Trace Collection System

Hongyang Zhao et al. proposed a system called Trace Collection System for friend recommendations in online social networking applications. In this system, authors mainly the focus on the friend recommendations by analyzing user's location by capturing few things. By capturing the location information of the users by analyzing data of GPS, Wi-Fi, and Bluetooth system will recommend the users in the social networks. Main motto of this concept is, people who are living or who are working in the same location they were interest to build the social circles with them. By capturing these data using the mobile phones, it will identify the social accounts of the users which are synchronize with the mobile devices. System will recommend the new friend requests after analyze the user ID of the online social accounts. In this following architecture demonstrated the Trace Collection system working procedure. Figure 1 explains the trace collection architecture, links that are established to user and device. In data management few sensors used for the effective communication to user and devices like wi-fi sensor Bluetooth sensor.
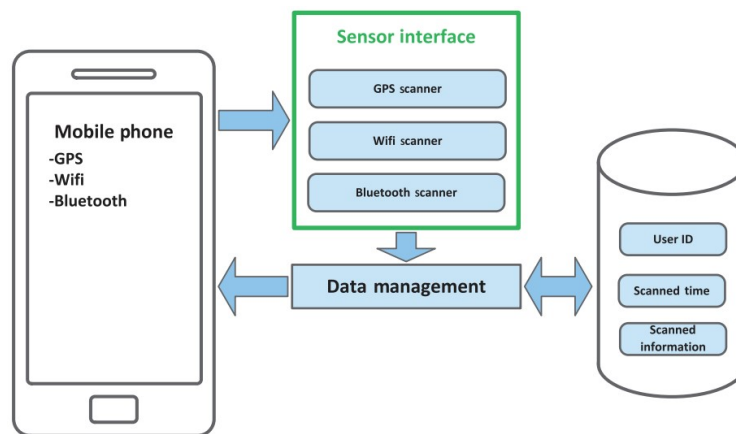

Figure 1. Trace collection system architecture

Friend Book

Wang et al. proposed architecture to suggest friends in online social media platforms. The system suggests friends based on the document search in the mobile devices. Generally, based on the user documents usages like a websites search, web pages accessing system will collect data of the documents and analyze the topics of the documents to identify the lifestyles of the users. For identification of topics in the life documents, system apply word frequencies of the document then it applies the LDA (Latent Dirichlet Allocation) algorithm to identify the topics of the documents. By implementing the friend graph construction, it will identify the lifestyles of the users. These procedures will synchronize with mobile social networking apps then system will recommend the users which have same lifestyle. In this following architecture demonstrated the Friend book system working procedure. Figure 2 it displays the architecture of the internal working of FRS that links with many features like life documents, friend query, user feedback.
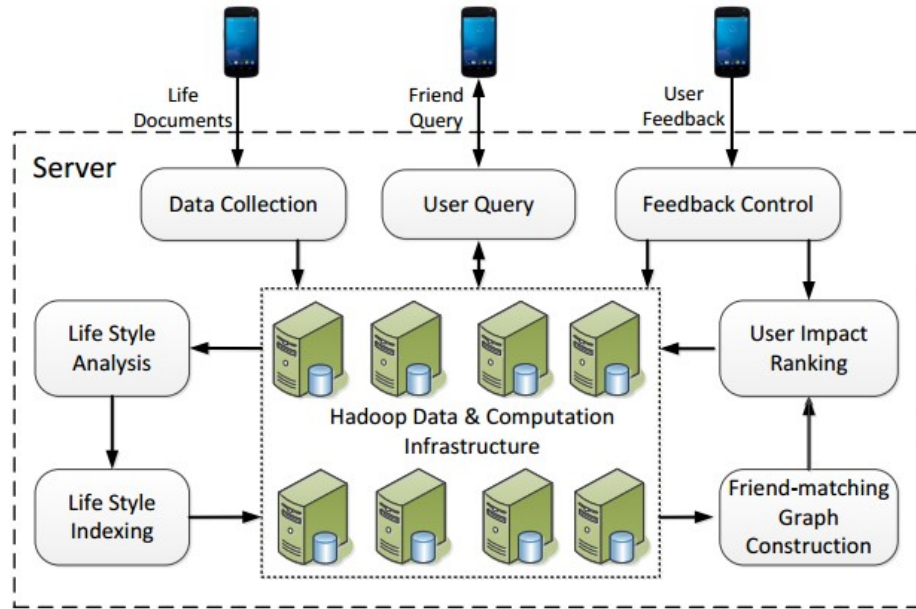
Figure 2. Friend Book Architecture

Match Maker System

Bian et al. proposed a system to suggest friends in a given online social media platform considering profile as attribute. This system monitors the user profile attributes like the list of shows by the users and list of the likes and dislikes of the shows etc. This system designed with a Matchmaker prototype in the Android platform to identify the personality matching by analyzing TV shows watching and characters liked by the users. Based on these parameters system analyze the similarity between the users and recommend the uses for building the social circle.

Linguistic Approach

Ch. Xu and J. Zhang designed a spam detection model with linguistic-based approaches but taken important and noticeable features. Here in this system, taken features are Product-based Sentiment, Brand-based Sentiment Deviation, and Brand-based Similar Text. By considering these linguistic features in this system it will identify the spam reviews.

II.LITERATURE REVIEW

After surveying the various relevant papers, the following inferences are drawn from different authors of different concepts

References	Authors	Concept	Disadvantages
1	Zhao, H., Zhou, H., Yuan, C., Huang, Y., & Chen, J.	**Trace Collection System** Based on the locations, system identifies the friend recommendations.	Not focus on the users personality behavior.
2	Z. Wang, J. Liao, Q. Cao, H. Qi and Z. Wang	**Friend Book:** Identify the personality of the users, by analyzing the Life documents.	Data collection is tough task and heavy computation required.
3	L. Bian and H. Holtzman	**Match Maker System** Based on the Profile attributes like Watching TV shows and characters identifies the personalities.	By analyzing the entertainment programs data to identifies the personality is unethical.
4	Ch. Xu and J. Zhang	**Linguistic Approach** Taken features of Product-based Sentiment, Brand-based Sentiment Deviation, and Brand-based Similar Text to identify the spam reviews.	Depends on only linguistic based approach.

5	Weiwei Zhang, Fangai Liu	**Probability matrix factorization** Taken features of trust relations on the social networks with different characteristics	Depend on only probability matrix factorization approach.
6	Daomeng Xu,Lu Jiang	**Topical Attention** Based the personalities the matching profiles are taken for the recommendation system in social networks.	Works with RNN only.
7	Peng M, Zeng G, Sun Z, Huang J, Wang H, & Tian G	**ARSM** This feature novel matrix factorization algorithm based on user's interest's applications permissions and functionalities.	This model Doesn't collaborate with PMF models.
8	Sarwar B., Karypis G., Konstan J., & Riedl J.	**Conventional CF** The search of neighbors are among a large uses population of potential neighbors it explores the relationship right in the first and rather than relationship between the users.	Considers only First and last relationship of the users.
9	Linden G, Smith B, York J.	**Information filtering, Information filters** It is used for Item-to-item Based on the Profile attributes like shopping, searching and characters identifies the personalities	-
10	lmongui, H. G., Mansour, R., Morsy, H., Khater, S., El-Sharkasy, A., & Ibrahim, R. TRUPI	**Clustering algorithms, Filtering algorithms** This features the users' personal interests, International Conference on Intelligent Text Processing	-
11	Davidson, J., Livingston, B., Sampath, D., Liebald, B., Liu, J., & Nandy, P., et al	**Information filters, Filtering algorithms** This features the users' personal interests' videos and recommendations on profile personalities.	-

This work only depends on the linguistic approaches for fraud detection. In the first feature, spammers post the same kind of sentiment similar on the products. For example, posting multiple reviews with a positive opinion. In the second feature Brand-based Sentiment Deviation, spammers keep posting the same similar opinion posting based on the brand not based on the product. For example, posting multiple negative reviews w.r.to brand. Last feature and Brand-based Similar Text, By considering these linguistic features in this system it will identify the spam reviews [1] to [4] and [9].

III. PROPOSED METHOD

In this proposed system, we came forward with architecture of Friend Recommendation System, which recommends the friends based on the calculation of the similar personality. Here we are considering the ratings of the users to calculating the personality similarity of the users. We can take any domain of the opinion data; in this project we are considering the E-commerce product related opinions are we are considering. Based on the harmony ratings of the users we define similarities of the users. To calculating the similar behaviors of the users we are implementing the Collaborative Filtering algorithm for the detection of harmony ratings. In our system Friend Recommendation System, we are considering the opinions main aspect to friend recommendation. In other hand it is very important to calculate the spam reviews and ratings, so that we also survey on the spam detection of the opinions.

In our proposed architecture we have three main modules. Those are Users, System, and Admin. In our proposed methodology users are the end-users and social users of our system. Here Users build a social circle with their friends and share data with their friends. In our system, users are E-commerce users also, in our system they can see the list of products uploaded by the admin and perform e-commerce operations like purchase products, post reviews, and ratings of the products, etc. In our system, users get friend recommendations based on the harmony of the ratings. This means taking the similarity of ratings between the users, the proposed system (Friend Recommendation System) will make recommendations to the users. In our system, Admin is the data owner of our application. Admin will add the products and product information like details of the product like name, cost, images, etc. In our system, based on the user's ratings and reviews CF (Collaborative Algorithm) will apply to calculate the similar personalities of the users. In addition, the system will also identify the review spam with three models. All these modules discuss in this chatter. In the following

figure, we describe our system clearly. Figure 3 shows the architecture of the whole system the link that are established to user and to the server.

Access Data

For demonstrating our system, we have taken E-commerce product reviews and ratings. For calculating the harmony of the ratings, we can take any data of ratings and reviews, but we have taken E-commerce data to more practicality. This access data feature of the users is a short form of total e-commerce operations by the users. In our system, users can see the product list and view data of product details, reviews, and ratings of the products. Users can post reviews and ratings.

Collect Users Feedback

Based on the products users can post reviews and ratings in our system. This feedback information is input data of the system to identify the friend recommendations. While posting reviews and ratings, systems will apply spam models to identify the spam reviews. Based on these reviews and ratings, system will calculate the similarity between the users and users will get Top-n recommendation based on the harmony ratings of the products.

Spam Features

In our proposed system we design the system for the prediction of friend recommendations to the users, for that we are taking reviews and ratings as input data. So, it is a high recommendation to apply the spam detection model for identifying spam reviews. For identification of the spam reviews, we used linguistic feature, behavior feature, and rating feature. Total three features we consider to detect the spam data. Those are,

1. Content Similarity (Linguistic Feature)
2. Rate Deviation (Rating Feature)
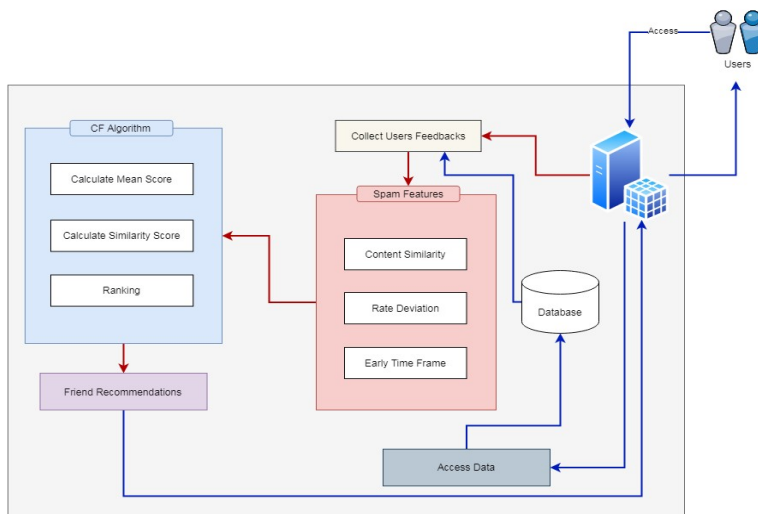3. Early Time Frame (Behavior Feature)

Figure 3. Friend Recommendation System Architecture

Content Similarity

This is a linguistic feature, which means this future completely depends on the text of the review. Generally, spammers are paid to write a bulk of reviews in a period of time, so they use to write similar kinds of text positively or negatively. By comparing the review text we can identify the spam reviews. In this procedure, if any two reviews of a user have similar text (Threshold $\phi=0.5$) we consider it to be spam. More than the threshold value of similarity is there for multiple reviews then the system considers it to be spam. For string comparison, we have used the Levenshtein Distance algorithm.

Rate Deviation

In this rating feature, we compare the ratings of users to a product with mean rating. If anyone post rating of the product then we compare with mean rating of the product, means average rating of the product. If any rating deviates abnormally then we consider being spam. To identify the rate deviation, we use following steps. Here ϕ is threshold value [0, 1], we consider to be 0.5.

$$1 - \left| r - avg(r\text{-}(p)) \right| / 4 = \psi$$
$$if \ \psi \geq \phi, then \ legitimate$$
$$else \ spam$$

Early Time Frame

This is based on context of the review data like dates of the review posting, so this is considered to be behavioral spam feature. Spammers need to write multiple reviews for a product, by taking this functionality we are considering the reviews who post in a short time to be spam. μ_i is a date of the post when the first review is done and μ_j is date of the post when last review is done.

$$\mu_i - \mu_j \in p$$
$$\text{if } (1 - (\mu_i - \mu_j)/7) > \phi, \text{ then spam} \quad \text{else legitimate}$$

Use Case Diagram

Figure 4. Use Case Diagram

Figure 4 shows a sequence diagram is a graphical view of a scenario that shows object interaction in a time-based sequence what happens first, what happens next. Sequence diagrams establish the roles of objects and help provide essential information to determine class responsibilities and interfaces.There are two main differences between sequence and collaboration diagrams: sequence diagrams show time-based object interaction while collaboration diagrams show how objects associate with each other. A sequence diagram has two dimensions: typically, vertical placement represents time and horizontal placement represents different objects.

CF Algorithm
Collaborative filtering algorithm is a popular to identify the likeminded people, using this algorithm we identify the similarity rating behavior by analyzing pattern of the ratings. Main important feature of this algorithm is it's considering the mean rating of the users with product ratings. For example, if user u has 1.5 is a mean rating, then user posts 3 to 3.5 rating for positive opinion. Likewise by considering the mean rating CF algorithm calculate the similarity of the two users a, b.
Let,
- $r_{a,p}$ is rate of the product p given by the user a
- $r_{b,p}$ is rate of the product p given by the user b
- $r`_a$ is a mean rating of the user a
- $r`_b$ is a mean rating of the user b

$$sim(a,b) = \frac{\sum_{p \in P}(r_{a,p} - \bar{r}_a)(r_{b,p} - \bar{r}_b)}{\sqrt{\sum_{p \in P}(r_{a,p} - \bar{r}_a)^2}\sqrt{\sum_{p \in P}(r_{b,p} - \bar{r}_b)^2}}$$

Proposed System Model

In this chapter we discuss about Friend Recommendation System and system flow of execution of the actors and operations with help of UML diagrams.

Requirement Analysis
Functional Requirements

Signup	This functionality describe that user should signup to get the login access. User can should provide basic details like name, email, contact number, profile pic, etc. while registration. After the user registration user can login to user portal.
Login	This functionality describes that admin and user should login for accessing their portal. User should register before login.
Search Users	In this functionality user can search the other users by providing the name or email id for sending friend requests to build the social circle.
Send Friend Request	In this functionality user can send friend requests to the known persons by after getting the results by the 'UC_03_Search_Users' to build the social circle.
Respond to Request	This functionality user will get the friend requests send by the other users and user can see and response to that request. User can accept the request or reject the requests.
Share Post	In this functionality, users can share something like image or text to their social circle. Users can see the posts of friends in their wall page.
Submit Rating	This functionality describe, users can submit their feedback on the product in terms of the raying and review. Users can see the ratings and review of the products.
Get Friend Recommendations	In this functionality, users will get the friend recommendations based on the harmony of the ratings after applying the CF algorithm.
Spam Detection	This functionality describe, System will spam detection model for identifying spam reviews while uploading the feedback by the users. Total three features we consider to detect the spam data. Those are, Content Similarity, Rate Deviation, and Early Time Frame.
Collaborating Filtering on Ratings	In this functionality, based on the user's ratings and reviews CF (Collaborative Algorithm) will apply by the system to calculate the similar personalities of the users

User characteristics

Admin	Admin is the data owner of our application. Admin will add the products and product information like details of the product like name, cost, images, etc. Admin has access to see the users data like profiles and feedback information.
User	Here users can builds social circle with their friends and share data with their friends. In our system, users are E-commerce users also, and can see the list of products uploaded by the admin and perform e-commerce operations like purchase products, post reviews, and ratings of the products, etc. Users get friend recommendations based on the harmony of the ratings.

Domain Requirements
We are developing a Java based Web application and we can develop many types of applications using Java. For executing J2EE framework we require an application server that can execute java code. Apache Tomcat is open source web server for executing the java web based applications. Application generates the data and we need to store in a medium. Mysql is open source database server we can store the data by using JDBC help.

User Requirements
- Any Update Processer
- Min 4 GB Ram
- Min 100 GB Hard Disk
- Windows Operating System
- Advanced Browser

Non-Functional Requirements
Product requirements section typically describes all the features, abilities and traits of a product that it possesses. To put it in layman terms, any function or constraint that the product is required to have to please the customer. The below is description of the same. Proposed model has proven to be three times faster than the current existing models that are

deployed. The product is absolutely reliable. At any given point of time, there isn't any requirement of third-party software that our model needs to depend on and the software operation is done through open-source programs. High-level programming language Java is used as Software portability in our application Development same software in different environments. This application can use by different persons like Users and Admin.

Organizational Requirements
Implementation Requirements
- Data Collection Phase
- Friend Recommendations

System Requirements
Hardware Requirements:
For making the project hardware processor of Intel Core *i7 Processor with* 16GB RAM and 256 GB hard disk and Only (64-bit) Version *has* been used. These specifications are subjective to a particular user and is not necessary that one should also have the same.

Software Requirements
Windows Family Operating System is recommended with Java and J2EE Technology, Web Technologies like Html, JavaScript, CSS and Apache Tomcat 7.0/8.0 web server has chosen. My SQL 5.5 or Higher Database and JDK 1.8 or Higher is used for this model.

IV. RESULTS AND DISCUSSION

For executing our project, we need to start Apache tomcat server. To start the tomcat server, we need to the Apache Tomcat installation folder and find the Bin folder in the installation path. From the Bin folder we can find the 'Tomcat8w.exe' file. Using that file, we can start the server. In the following figure we can see how we start server. The Apache Tomcat displays the properties of the application where can we start and stop to run the local host. It also establishes the link to local host and database as shown in figure 5.
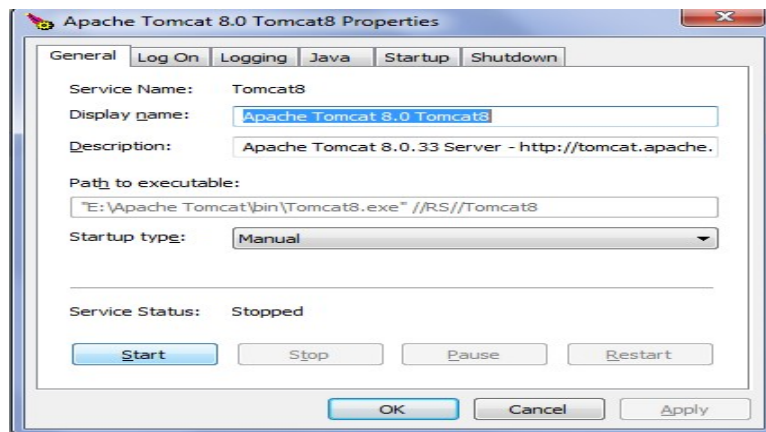
Figure 5. Start Server

After starting the server, we can access the project from the home page by entering following URL in the browser. http://localhost:post_no/webappname/
From this particular webpage (home page) one can guide himself towards user registration section, user login and also admin login section. After clicking User page, user can see the result page shown in figure 6 (a). From this page user can sign up and login. From the home page admin can navigate to the login page. By providing admin user id and password admin can navigate to the admin portal. After successful login, authorized person can have access for admin operations. Admin can add the main menu for the products. Figure 6 shows the step by step execution results of proposed algorithm to get friend recommendation based on reviews and user ratings.

In this project to store the data we have used SQLyog. The data will be collected and modified by the admin. The Apache TomCat8 establishes the network to the local host in the system to access the database. Figure 6(a) shows the home page which have access to admin whereas figure 6(b) shows the sign up and login in options where new users can register or existing user can login. Figure 6(c) shows the admin user page where the data can be observed and instructed the user to

choose to add product based on data provided, add sub category, product information, view ratings, view products, purchase history, order details and feedbacks. In figure 6(d), it has shown that the data in admin user pages that is which displays the categories of particular products with necessary relevant data. Figure 6(e) displays the products that should be given in respective categories like in mobile category different mobile models are listed for display. In figure 6(f) it displays each products buyer's information with his Id number and given rating of the products. Figure 6(g) shows the information of the product and every detail that is provided by the admin which is saved in the database SQLyog which can be modified only by the admin. Figure 6(h) shows the ratings that are given by the users to check that if the ID is genuine or fake, it will be reviewed by the non-spammer algorithms used in this model. Figure 6(i) displays the products to the users who wants to buy whereas figure 6(j) displayed the purchases history of the product by the user.

In figure 6(k), it displayed the order details of the product for user visibility purpose and figure 6(l) shows the products owned by the users can give a feedback to the admin stating that how the products perform or any damaged shipping or other complaints that can be provided in the feedback. Figure 7 displays the Friend Recommendations according to the user personals information as per the products purchased, rating and reviewing with help of non-Spammer algorithms. Also, we can detect the fake profile users that are considered as spammers and by eliminating them, friends are recommended to users. So, user can send Friend requests.
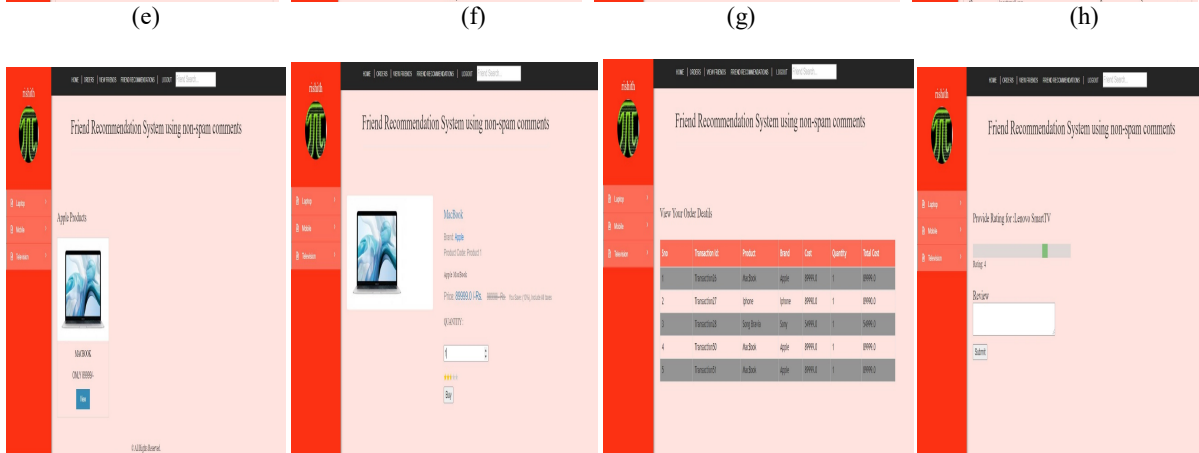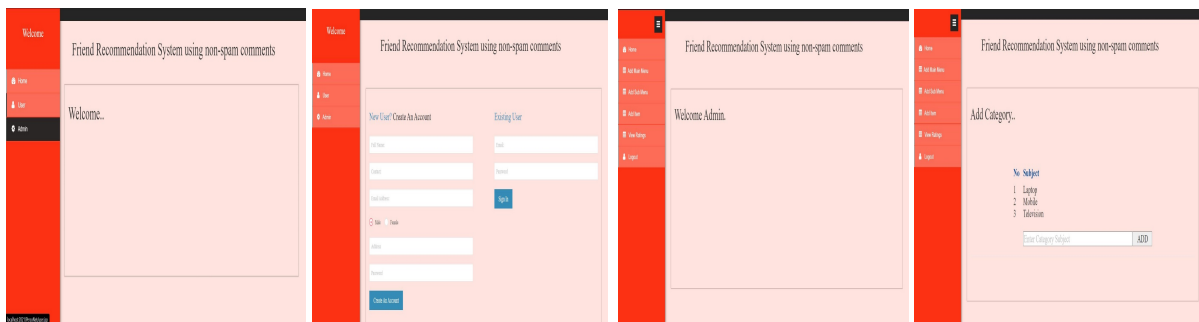

(a)   (b)   (c)   (d)

(e)   (f)   (g)   (h)

(i)   (j)   (k)   (l)

Figure 6. (a) Home page (b) Signup and Login page (c) Admin Home Page (d) Add Product main menu (e) Add Product Sub Category Names (f)Add Products (g)View Product Information (h) View Ratings . (i) View Products (j)Purchase Products (k) Order Details (l) Feedback Page
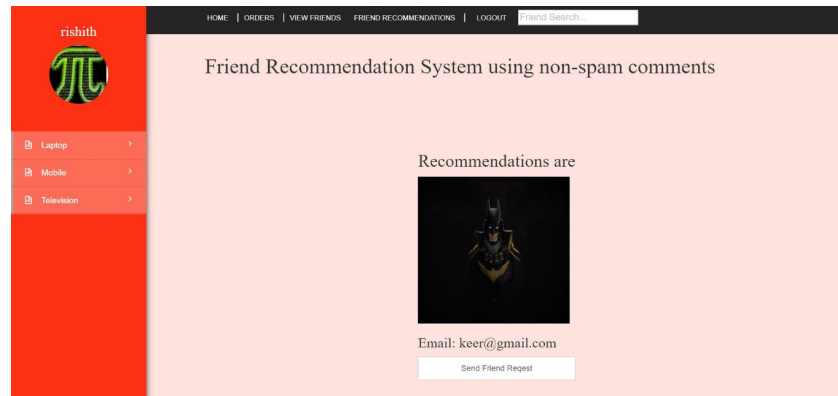

Figure 7. Friend Recommendation

CONCLUSION

In the past decade, social networking applications are the most commonly used services that provide an online platform for maintaining social relations and build social circles with others who are known to have similar personalities, careen similarities, etc. Most of the social network-based applications provide the facilities to create and update the social circle manually, which means users should select or search profiles of the users for building the social circle based on their interests. The current social networking applications recommend friends based on mutual friends, contacts match, and other aspects. Many surveys proposed friend recommendations based on the profile attributes like age, location, interests, etc. In our system, we came up with Friend Recommendation System, which makes recommendations based on the harmony ratings of the users. We design our system by implementing the Collaborative Filtering algorithm for the detection of harmony ratings. Our system should be able to identify the spam reviews as our input consists of ratings and reviews, so it is very important to notice the concepts of the spam review detection. Based on the results our system recommends friends to the users.

Currently we are using machine learning to develop this type of friend recommendation system. This could be accurate but it could take time to search and recommend friends. To further this development and make it more efficient we can develop this even further by using NFC (near field communication) technology. We can allow users to add friends using NFC when they are nearby which can make them add friends very quickly and without any hassle. This could help users to save time to add friends instead of the system recommending friends. NFC is widely popular and is used in many smartphones which are manufactured in this recent time. So, this will not add any cost to the user as it is readily available**.**

REFERENCES

[1] S. KEMP. (2018). Number of Social Media Users Passes 3 Billion With no Signs of slowing. https://thenextweb.com/ contributors/2017/08/07/number-social-media-user%s-passes-3-billionno-signs-slowing/

[2] W. Youyou, D. Stillwell, H. A. Schwartz, and M. Kosinski, "Birds of a feather do flock together: Behavior-based personality-assessment method reveals personality similarity among couples and friends," Psychol. Sci., vol. 28, no. 3, pp. 276–284, 2017.

[3] Zhao, H., Zhou, H., Yuan, C., Huang, Y., & Chen, J. (2015). *Social Discovery: Exploring the Correlation Among Three-Dimensional Social Relationships. IEEE Transactions on Computational Social Systems, 2(3), 77–87.* doi:10.1109/tcss.2016.2517092.

[4] Z. Wang, J. Liao, Q. Cao, H. Qi and Z. Wang, "Friendbook: A Semantic-Based Friend Recommendation System for Social Networks," in IEEE Transactions on Mobile Computing, vol. 14, no. 3, pp. 538-551, 1 March 2015, doi: 10.1109/TMC.2014.2322373.

[5] L. Bian and H. Holtzman. Online friend recommendation through personality matching and collaborative filtering. Proc. of UBICOMM, pages 230-235, 2011

[6] Xu, Chang. (2013). Detecting Collusive Spammers in Online Review Communities. 10.1145/2513166.2513176.

[7] Zhang W, Liu F, (2019) Recommendation system in social networks using probabilistic matrix factorization. PLoS ONE 14(10): e0223967. https://doi.org/10.1371/journal.pone.0223967

[8] Xu D, Jiang L (2019) Recommendation system in social networks with topical attention. PLoS ONE 22(13): e0223967. https://doi.org/10.1371/journal.pone.02235767

[9] Peng M, Zeng G, Sun Z, Huang J, Wang H, & Tian G. Personalized app recommendation based on app permissions. World Wide Web.2018;21(1):89–104.
[10] Sarwar B., Karypis G., Konstan J., & Riedl J. Item-based collaborative filtering recommendation algorithms. Www. 2001;1: 285–295.
[11] Linden G, Smith B, York J. Amazon. com recommendations: Item-to-item collaborative filtering. IEEE Internet computing. 2003;(1): 76–80.