SIMILAR SOCIAL TRAITS IDENTIFICATION SYSTEM

PROJECT REFERENCE NO.: 39S_BE_1967

COLLEGE : ATRIA INSTITUTE OF TECHNOLOGY, BENGALURU
BRANCH : DEPARTMENT OF COMPUTER SCIENCE ENGINEERING
GUIDE : MRS.HEMALATHA K N
STUDENTS : MS. ANJANA RAO N
           MS. POOJITHA S PRASAD
           MS. ARCHANA M V

KEYWORDS:
Friend recommendation, mobile sensing, social networks, life style.

INTRODUCTION:
The existing social networking websites caters friend proposition to users based on pre-existing user relationships or that determined by the geographical distances. The results obtained from the system, may not be the most efficient to reflect users preferences in picking friends. As a substitute, certain mechanisms which make recommendations based on Lifestyles and Interests can be used. Where the lifestyles refer to activities accomplished by an individual on a daily basis.

Buddy-Finder, is a recommendation system, which recommends friends to users based on their interests and social traits. The Smartphones such as iPhone or Android –based phones are equipped with a rich set of embedded sensors, which serves as an ideal platform for sensing daily routines from which people’s lifestyles could be discovered. The user’s interests are collected from the web portal and the users lifestyles are extracted using LDA algorithm from the activities captured from the smartphones. Upon receiving a request, Buddy-Finder returns a list of people with highest similarity scores.

OBJECTIVE:
The objective of project is to recommend friendship in social networking site by integrated activity learning. When people with similar interests are recommended as friends, then user interest in activities enhances as he can keep in touch with interested friends and share knowledge also community building can be enhanced with this new approach of friendship recommendation. It is easy for people to get friends for jogging, shopping etc in their close quarters. People with similar interests can establish a virtual community and share knowledge using the social network platform with our approach.
METHODOLOGY:

DESCRIPTION OF MODULES
This section describes each of the above modules in brief.

- **Module: Login Or Register**
  This module contains login or registration in order to register the user with Buddy Finder application, during the registration the users interests will be taken.

- **Module: Authenticate Users**
  This module compares the entered Username and Password with the respective records among the database entries. If a match is found, then redirects the user to his/her profile page. Else, the user is redirected to the registration page.

- **Module: Collect_Data_From_Sensors and Interests from users**
  1. Each smartphone records data of its user from the sensors such as accelerometer, gyroscope and GPS information.
  2. The recorded data is then converted in to activity sequence and is sent to the server for further processing in the format \{acc_x, acc_y, acc_z, gyr_x, gyr_y, gyr_z\}.
  3. Also when the user is registering he/she has to select their interests.

- **Module: Data Collection and Pre-processing**
  This module collects the data sent from the client side. The data collected will be in format \{acc_x, acc_y, acc_z, gyr_x, gyr_y, gyr_z\}. The collected raw data is further pre processed to remove outliers. Median filtering technique is used for outlier detection and removal. An unsupervised learning technique is applied on the preprocessed data to form clusters known as Kmeans clustering algorithm. The resulting clusters forms a list of activities carried out by a user, where each cluster representing an activity.
• **Module: Activity Recognition**

Now that the k clusters are formed, each represents an activity, lifestyles are further extracted from these activities using LDA algorithm. Once the similarity is calculated, the user who has highest similarity is suggested as a friend.

**RESULTS:**

1. Computation of centroid values to be given to Client:
   a. A huge dataset, consisting of raw sensor values in the form \((\text{acc}_x, \text{acc}_y, \text{acc}_z, \text{gyr}_x, \text{gyr}_y, \text{gyr}_z)\) are given to server.
   b. The raw data are further filtered using median filtering for outlier removal.
   c. K Means algorithm is applied on the filtered data with pre defined value of K as 10, where K represents K activities (clusters).
   d. Centroid values obtained from K-means clustering, for K clusters are stored in a file.
   e. The resulting file with centroid values are distributed to all the users, who are registered with the application.

2. Friend computation:
   a. The user has to configure the server by providing server IP address.
   b. The user has to next register with the app. After which the user can start using the buddy-finder service.
   c. Instead of raw data the activity sequence are sent to the server. Thus reducing the overhead of pre-processing computation at the server end.

**CONCLUSION:**

Unlike the current friend recommendation schemes, which depend on the preexisting social relationships and geographical information, “Lifestyle based friend recommendation system” is a scheme where the friend suggestions are provided based on users daily activities. The Buddy Finder mobile application captures the user’s daily activities from their smart phones and suggests friends to users if they share identical lifestyles. The proposed scheme is implemented as a mobile App on the android devices. And small scale experiments are performed to evaluate the system performance. This report provides a detailed description of the design, implementation and analysis of the results obtained by the experiments conducted on the system.

**FUTURE WORK:**

Beyond the present model, the future work can be a three-fold. First, the Buddy Finder application can be integrated into current real social networking services such as Facebook, twitter and so on. Second to conduct large-scale experiments and evaluate the system to be scalable to large systems. The similarity threshold used in the experiment is fixed. Experiments could be conducted with varying threshold and a comparison on the results obtained with varying threshold could be computed. Third, more sensors could be incorporated on mobile phones, into the systems and also make use of wearable sensors, so that the system could use more information for lifestyle extraction, which would improve the recommendation accuracy.