Abstract— Social Networking Service (SNS) is defined as the popular online platform which is useful to build social networks and social activities of people by sharing their real-life connections, career interests, and background activities. Most of the existing works are devoted to SNS recommender systems and CPSSs (Cyber-Physical-Social Systems) is difficult to analyze user’s preferences and behavior patterns. SNSs have attained a progressive growth with extensive data in recommendation systems that increase the complexity levels. In this paper, an extensive integration of cyber, physical and social spaces discussed to overcome the issues of the existing system like difficulty in sequential pattern mining, unavailability of group user in individual-centric recommender systems and low dimensionality rating data. Cyber-Physical-Social Systems (CPSSs) have become the simple paradigm of evolution in the information industry, through which traditional computer science will develop into cyber-physical-social computational science. Intelligent recommender systems are an important fundamental research topic in the CPSS field and the implementation of personalized and intelligent computing has great significance in CPSS development. This paper proposes a group-centric recommender system in the CPSS domain with activity-oriented group discovery, the revision of rating data which are analyzed in terms of accuracy, precision analysis, recall analysis and level analysis.

Keywords—Social Network Services, Cyber-Physical Social Systems, Group discovery, Intelligent recommender system.

1. INTRODUCTION

With the advancement in Social Network Services (SNS) with increasing amounts of data from physical, social and cyber spaces are generated and distributed rapidly [1]. Currently, the world involves social activities, gatherings and cyber resources, through multidimensional comprehensive systems that provide computing, communication, control, commerce, etc. Due to the innovation in mobile and sensing technologies, social data are easily collected as records of social events, which are analyzed to yield a better understanding of user’s daily activities, social circles, living habits, etc. Hence, the need of focus on developing community centric applications based on data analysis that integrates social, mobile and big data technologies, such as Cyber Physical Social Systems (CPSSs) becomes more necessary. Intelligent recommender systems are the significant technology and a typical class of applications for providing personalized and intelligent services in the CPSS paradigm. The issues in Volume, Variety, Velocity and Veracity, lead to the efficient analysis of the big data collected in CPSSs. In order to overcome the drawbacks of the existing system like Complexity in sequential pattern mining, unavailability of group user in individual-centric recommender systems and low dimensionality rating data, an effective proposed system is needed.

In the proposed system, a group centric intelligent recommender system named GroRec is introduced to integrate social, mobile, big data technologies and accurate recommendation services in CPSSs and decrease the complexity of conventional individual centric recommender systems. Hence, a quantitative method of assessing emotional offset, based on sentiment analysis, reviews and ratings are used to improve the objectivity of context data are described as in Fig. 1.
The remaining section of this paper is organized as follows: Section II reviews some of the existing works related recommendation system integrating social media and big data analysis. Section III provides the detailed description of the overall proposed system. Section IV presents the performance results of the proposed system. Finally, this paper is concluded in Section V.

2. RELATED WORK

This section presents some of the existing works related to recommendation systems which integrate social, mobile and big Data. Zhu, et al. [2] proposed a novel modeling and computing technique for geological information service system developed in complex data processing. Also, certain techniques of modeling the information service system and computing geological data by cyber-physical-social-thinking-modelling and computing geological information services system were listed. Salehi-Abari and Boutilier [3] introduced a preference-oriented social networks (POSNs) to capture the correlation of preference rankings between individuals interacting in social networks. They developed methods for group recommendation and missing references and clearly explained the way of support to the accurate inference of preference rankings. Hence, Preference Oriented Social Networks (PSON), a generative model was proposed to determine the similarity of the preference rankings of individuals.

Boratto and Carta [4] suggested a recommendation system to explore the rating prediction task with an objective to identify the effective method to forecast the ratings for a group which is automatically detected by the system. Chen, et al. [5] evaluated context-aware splitting approaches and differential context modelling algorithms. They examined predictive performance, and also explored the usage of emotions to discover how emotional features interact with those context-aware recommendation. Chen, et al. [6] introduced a novel group recommendation method via fusing the modified collaborative filtering methodology with the temporal factor in order to, solve the dynamics problem.

Sheth and Anantharam [7] Recommended Physical-Cyber-Social (PCS) computing, that implemented a human centric and holistic view of computing by analysing observations, knowledge, and experiences from physical, cyber, and social worlds. Hussein, et al. [8] proposed smart services framework in CPSS (called Dynamic Social Structure of Things, or DSSoT) boosts sociality and narrows down the contextual complexity based on situational awareness. An application using DSSoT, called Airport Dynamic Social, provided a proof of concept. DSSoT observed spatio-temporal situations based on user’s individual goals and other social aspects, structures relevant to social objects in a temporal network of interactions.


Yang, et al. [11] proposed a novel social CF method called TrustMF by the heuristic algorithm which affected the process of reviewing. A trusted model and a trustee model was introduced to map the users into same latent feature spaces with different implications.

Ning and Liu [12] explored an updated science and technology framework for the IoT and a novel cyber-physical-social-thinking (CPST) space was established by involving Internet of Thinking (IOTk). The newly introduced CPST involves distinct space dimensions for the IoT as listed below:

- Cyber space: Cyber space denotes the virtual, digital and abstraction to achieve the interconnections.
- Physical Space: Physical space involves the real world in linear dimensions.
• Social space: Social space is described as the intra-relationships with other associated things.

Li, et al. [13] suggested an effective privacy-preserving item-based collaborative filtering algorithm to obtain the recommendation accuracy and calculate the efficiency. Item similarities and correlations were incremented by computing a secure multi-party computation protocol. Wu, et al. [14] introduced a novel problem of explaining dynamic semantics to the mobility data with the help of contextual data. The purpose and interest of a particular user from his location added a benefit to the wide range of application in the world. They assumed two inputs namely,
• Location History
• Social Media Documents

Liu, et al. [15] proposed an algorithm to estimate the missing values in the tensors of visual data. The value might miss due to the acquisition process and algorithm worked even with a few of samples and it propagated structure to fill larger missing regions. The contribution of the paper is to prolong the matrix case to the tensor case by suggesting the initial definition of the trace norm for tensors and then by building a working algorithm. First, they proposed a definition for the tensor trace norm that generalizes the established definition of the matrix trace norm.

Pirasteh, et al. [16] introduced new weighting algorithm between the users to improve the results significantly by adding traditional similarity measures and transforming symmetric similarity to asymmetric similarity based on user comments on non-common items. The habit and effects of users were considered by their ratings and comments repetitively on common rated items. Experiments on datasets were implemented and compared to other similarity measures.

3. PROPOSED METHOD

This section presents the detailed description to decrease the complexity of conventional individual centric recommender systems. The overall process of the proposed system is shown in Fig.2, which includes the following stages:
• Group Discovery
• Sentiment Analysis
• Review Analysis
• Rating Analysis
• Recommendation System

Fig.2 Overall flow of the Proposed System

3.1 Group Discovery

In order to address the issue of the expansion of context data available in CPSS, a group-centric methodology was applied to decrease the computational complexity of the recommender system. Though, the conventional group discovery approach is focused on clustering of low-dimensional data and it is not applicable for CPSSs. In GroRec, the group discovery groups from high-dimensional data are focused on the similarity of user behavior. Group Discovery in GroRec comprises following steps namely,
• Tensor-based representation of user behavior
• Behavioral Similarity Analysis via Tucker Decomposition
• Group Discovery via clustering

The purpose of the group discovery is to mainly categorize the users based on behavioral similarity. In a CPSS, the available user behavior data have certain elements like time, user, place and behavior. A SUB model is defined as a data structure which uses a mathematical approach in methodologies. A tensor represents a multidimensional array with the suitable means of representing the SUB model in three-dimensional data. The group centric data fusion is calculated on the basis of behavior weight is given by

\[ W_b = \frac{a_b}{\sum_{i \in \text{group}} a_i} \in A' \]

Here, \( a_b \) is the element corresponding to behavior \( b \) in approximate tensor \( A' \) and \( \sum_{i \in \text{group}} a_i \) is the sum of all elements present in the group. Hence, through group-centric
data fusion, the behavior data of all members are integrated into an individual user.

3.2 Sentiment Analysis

Sentiment Analysis is defined as the process of task extracting the positive and negative feeling and opinion of users about the entities like brand or product in social web services, article blogs and social media. This analysis helps the people to perform decision making on various entities. It is the method of collecting the feedback on the products or services by improving the quality and detecting emotional changes. These lead to modeling sentiment comments, recognizing and matching relevant contexts in multiple entities and sarcasm or noise content filtering. This analyses the method of focusing the well-structured tag information provided by the most social network service providers.

In this module Sentiment analysis, users load the group discovery data from the database. Sentiment analysis is used to analyze the sentiment based on the user data denoting positive expression of the users. Finally, after completing the analysis, particular data is stored in the database as shown in Fig.4.

3.3 Review Analysis:

Based on the valuable information extracted from reviews, the topic modelling and opinion mining are advanced. Assessment of products, social and e-commerce media are encouraged to write reviews in the form of textual comments explain like or dislike, good or bad. The review elements are described in general ways namely,

- Frequent terms
- Review Topics
- Over-all opinions
- Review Emotions
- Contextual opinions

In this review analysis, users load the group discovery data from the database. Review Analysis is process of analyzing the review based on the user’s feedback content. After complete analysis, the data is stored in the database are depicted in Fig.5.
3.4 Rating Analysis:

In GroRec, the emotional offset is computed through sentiment and review analysis and used to revise the respective rating data in order to improve the user evaluation objective. The emotional offset is computed by non-supervisory learning method focusing on three methods like,

- Text Preprocessing
- Emotional offset computation
- Rating revision

The reviews contain text, preprocessed by word segmentation, morphological normalization and removal of stop words and punctual expressions. The emotional offset of each review is computed and based on this only each rating is revised.

In Rating Analysis module, users load the group discovery data from the database. Filter the data based the rating. Rating means user give rating based of the user opinion as in Fig.6.

The weight of the user interests and item features are $\alpha$ and $\beta$ where $A$ is considered as the latent-topic-mapping matrix of the latent factor-rating matrix

$$J = \sum_{g_i \in \text{train}} (R_{g_i} - P_g Q_f)^2$$
$$+ \alpha \sum_{g_i \in \text{train}} (G - P_g A^T)^2$$
$$+ \beta \sum_{g_i \in \text{train}} (I - Q_i A^T)^2$$
$$+ \lambda (||P||^2_F + ||Q||^2_F + ||A||^2_F)$$

$$Sim(g, g') = \varepsilon P_{rat} (g, g') + (1 - \varepsilon) P_{int} (g, g')$$

3.4 Recommendation System:

The significant idea of the collaborative filtering is to utilize the feedback of the individual and after concerning the feedback, a large number of users and feedback is available to determine the similar user. Based on the particular ratings for products, recommendation accuracy is decided. Identifying and preparing the list of similar comments from the users are used to predict the rating values of similar products as in Fig.7.

The recommendation system is used to load the sentiment, review, rating analysis from the database. The recommended item based on the user input using the collaborative filtering.

Fig.7 Recommendation Analysis

4. PERFORMANCE ANALYSIS

This section presents the performance results of the proposed Collaborative Filtering algorithm. The results are analyzed and evaluated in terms of

- Accuracy
- Precision Analysis
- Level Analysis
- Recall

Moreover, the proposed collaborative filtering technique is compared with the existing techniques for proving the better performance of the proposed system. From this analysis, it is proved that the proposed collaborative filtering gives the best results.

4.1 Precision Analysis

Precision is said to be the probability that a (casually designated) retrieved document occur in relevant. From the Fig. 8, it is well known that $x$–axis represents the levels and $y$–axis denotes the precision rate. The proposed system is less when compared with other techniques. MF shows 81% and GROREC shows 78% at maximum level 35. Especially, at maximum level, proposed work exhibits minimum precision rate with 3.70% improvement.
4.2 Level Analysis

The objective of level analysis is to precisely forecast the target class for each case in the data. It is classified into MF and CF analysis. From the Fig.9, it is well known that x-axis represents the CF stage in terms of percentage and y-axis denotes the MF level. The proposed system is high when compared with other techniques. MF shows 9% and GROREC shows 10% at maximum level 19. Especially, at maximum level, proposed work exhibits maximum MF level with 10% improvement.

Fig.9 Level Analysis (MF)

From the Fig.10, it is well known that x-axis represents the stages and y-axis denotes the CF levels. The proposed system offers high MF levels for high CF stages. CF shows 9 and GROREC shows 10 levels at maximum level 25. Especially, at maximum level, proposed work exhibits minimum precision rate with 10% improvement.

Fig.10 Level Analysis (ICF)

4.3 Recall

Recall is the probability that a (randomly nominated) relevant document is recovered in specific search. From the Fig. 11, it is well known that x-axis represents methodologies involved and y-axis denotes the recall in terms of percentage. MF shows 85% and GROREC shows 75% at maximum level 100. Especially, at maximum level, proposed work exhibits minimum precision rate with 11.76% improvement.

Fig.11 Recall Analysis

4.4 Accuracy

Accuracy of a filtering technique is termed as the capability of correctness in filter. An amount of a predictive model that replicates the proportionate number of times that the model is correct when it is applied to data.

\[
\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of all cases to be predicted}}
\]

From the Fig.12, it is well known that x-axis represents time variations involved and y-axis denotes the accuracy in terms of percentage. For time variation 1 proposed work
shows 88% and it shows 100% at time variation 2. Specifically, at maximum level, proposed work exhibits maximum accuracy rate with 12% improvement.

**Fig. 12 Accuracy analysis**

5. CONCLUSION AND FUTURE WORK

This paper proposed a group centric intelligent recommendation system named GroRec, which integrates social, mobile and big data technologies to provide effective, objective and accurate recommendation services in CPSSs. Specifically, a group-centric approach based on the behavioral similarity among users is proposed to decrease the complexity of conventional individual centric recommender systems and a method of quantifying emotional offset based on sentiment analysis is proposed to revise user ratings for improved objectivity of rating data. A comprehensive approach based on three aspects like ratings, interests and social relationships are used to comprehensively extract group preferences. Furthermore, experiments verify that the performance of the proposed GroRec system is superior to that of the conventional CF-based and MF-based approaches. However, the proposed scheme can still be further improved, the KNN-based group discovery module has limited processing power. Therefore, the future work attempts to develop a deep learning based approach to improve the efficiency and accuracy of group discovery.

REFERENCES


