Data-Driven Diffusion Recommendation in Online Social Networks for the Internet of People

Diyawu Mumin[1, 2], Lei-Lei Shi[1, 2], Lu Liu[3], John Panneerselvam[3]
[1]School of Computer Science and Telecommunication Engineering, Jiangsu University, China
[2]Jiangsu Key Laboratory of Security Technology for Industrial Cyberspace, Jiangsu University, China
[3]School of Informatics, University of Leicester, UK
Corresponding author: Lu Liu, Email: l.liu@leicester.ac.uk

Abstract – Recommendation systems are gaining popularity with the proliferation of the Internet of People (IoP). The popularity and use of online social networks facilitate integrating these social relationships with recommender systems under a single framework of IoP. This paper proposes a new approach for item recommendation based on the diffusion method that combines user relationships in social networks with user-item relationships derived from the IoP. Especially, a resource redistribution process is explored in the user–object network that gives mass diffusion a higher recommendation accuracy and Heat Conduct a greater diversity by considering the social degree of users whilst calculating the user degree in the network. A tuning parameter is introduced to adjust the weight of resources that the objects finally receives from users based on their social relationships. Finally, extensive experiments conducted on the real-world datasets which contain friendship relationships, demonstrate the efficiencies of our proposed method in achieving notable performance improvements in terms of the recommendation accuracy, service diversity, and practical dependability.

Keywords – Recommender Systems; Social Networks; Item Recommendation; Internet of People

1. INTRODUCTION

Given the proliferation of the Internet of People (IoP) and the recent improvements in the storage capacities of the IoT devices [1], it has now become possible for users to generate and store different types of information on a daily base [2]. Storing such daily information avails great benefits as it is possible to retrieve any useful information in IoP by using an efficient resource discovery and recommendation system [3, 4]. However, an appropriate description of keywords is crucial during a search process since irrelevant keywords might mislead the search process. To assist such search scenarios, recommendation systems have been introduced to retrieve and recommend appropriate services to user queries in IoP.

Recommendation systems have been a popular research topic which has attracted attention worldwide [5, 6]. Existing research on recommendation systems exploits emerging technologies such as, cloud computing [7], software engineering [8] and service-oriented computing [9], most of such approaches rely on QoS based strategies [9-11]. Major researches in recommendation algorithms can be divided into three main categories such as Collaborative Filtering (CF) [12], Content-Based Filtering [13], and Hybrid Recommendation approach [14]. Recommendation systems provide personalized services to both individual users and a variety of web-based applications such as electronic commerce [15-17], electronic government [18], electronic education [19], and electronic health services [20]. Content-based approaches involve recommending objects that are similar to the objects previously liked by corresponding users, depending on computing correlations between relevant objects [3]. To control the system from overspecializing on certain objects, collaborative filtering has been developed believing that similar users usually have similar interests. Furthermore, the hybrid recommendation method exploits the merits of both collaborative and content-based approaches by combining them to avail personalized services to users, thereby overcoming the issues of cold start exists in many existing methods, and to achieve diversity in service recommendation [9]. However, balancing the emphasis given to the contents and collaboration whilst identifying services in the hybrid approach is still an open issue.

There have been further studies in recommendation systems research based on complex networks leading to the development of diffusion-based recommendation algorithms. Complex networks use the network structures rather than the rating matrix to model user-object bipartite graphs, making diffusion-based algorithms rely on unary data without ratings [21-23]. Mass Diffusion (MD) [24] and Heat Conduct (HC) [25] are two of the earliest recommendation algorithms proposed based on diffusion. While MD is based on physical dynamics such as random walk, with a higher accuracy, HC is based on heat conduct, and has a higher diversity. Much research has been done to resolve the accuracy-diversity dilemma in recommendation systems. To improve the accuracy of the HC model, [26] considered bias heat conduction, and weighted heat conduction. Also, there has been further research [27-29] on improving the diversity of the mass diffusion models by removing redundant correlation, using clustering, and, by using multichannel diffusion. Taking the merits of MD and HC into consideration, [27, 30] in their various research came out with different hybrid diffusion algorithms.

Some researchers treat the ratings differently in network-based recommendation algorithms. While Da-Cheng et al. [30] treat ratings equal in their research, others discard negative ratings during preprocessing [21, 31, 32], and some proved the important contribution of negative ratings in either dense or sparse datasets [33, 34]. Many researchers drop negative ratings due to their effect on the computational complexity of the algorithms, this is because more variables have to be introduced in assigning resources and weights for the negative ratings [34].

Most of the aforementioned recommendation methodologies are designed predominantly based on user ratings without considering user relationships. Nevertheless, user preference mostly depends on social relationships among the existing users in online social networks (OSNs) [35, 36]. Also, user purchase behaviors are influenced by their social connections [37].

Therefore, some previous studies have established that users in the same group share similar properties [38-40]. For example, Shen et al. [41] used a user-friendship and user collaborative filtering approach to proposed a friendship-based method, Guo et al. [42] proposed a matrix-factorization method based on the integration of multiple user relationships. Wang et al. [24] incorporate users' explicit and implicit trust and proposed a resource-allocation method on a tripartite graph. This is to insist that user opinions pose significant influence whilst choosing objects in real life [7, 43]. User interest is dynamic and changes over time, thus, both the user interests and social connections should be incorporated into recommendation systems for availing relevant services. However, there have not been many studies on integrating user relationships into diffusion-based methods [44], calling for further studies on developing better methods based on the network-based framework.

In this paper, a friendship-based diffusion recommendation method is proposed, named, FDR, by introducing the user-user relations into the resource allocation process of network diffusion using the inverse log function. Specifically, the proposed method exploits user relationships during the recommendation process, by scaling the amount of resources that a user receives using a tuning parameter before it is redistributed back to the objects at the final stage of diffusion. An optimal value of the parameter gives a better result as measured by the recommendation accuracy across the different datasets. Several experiments are conducted using two different real-world coupled datasets, namely Epinions [45] and Friendfeed [46], to measure the accuracy of the recommended services to users based on user's friendly relationships. The experimental results show that exploiting social information such as a user's friendship relationship enhances objects diversity and also improves the accuracy in object recommendation. The main contributions of this paper are summarized as follows.

1) An enhanced data-driven recommendation algorithm developed based on user-object diffusion in online social networks, named FDR (Friendship-based Diffusion Recommendation) method. It considers both the user-object distribution and the relationships among friends [56] in the social network to redistribute resources in the user–object matrix. A tuning parameter is introduced to adjust the weight of resources that the objects finally receives from users based on the users' social relationships, which addresses the issues of data sparsity, thereby improving the recommendation accuracy and diversity.
2) Objects yet to be seen or rated by users are extracted and ranked in descending order based on their resource scores through the diffusion process, with the top-K objects recommended to the target user.
3) Several experiments conducted using two real-world datasets [56], Epinions and Friendfeed, demonstrate the effectiveness of the proposed model in terms of the achieved diversity and accuracy of the top-K recommended items.

The remainder of the paper is structured in the following order: Prior studies related to recommendation systems are discussed in Section 2, Section 3 introduces our proposed model along with the compared baseline models. In Section 4 we present a discussion on our experiments while Section 5 concludes our work and outline our future research directions.

2. RELATED WORK

Recommender systems are on the rise due to the success of social networks, the features of OSNs also gears toward improving the recommendation accuracy [47] by taking strategic advantage of the friendship relationships existing between the users. This section reviews the related works in recommendation systems of various categories.

One of the main approaches to item recommendation in online social networks is the Memory-based approach. Memory-based social recommender systems exploit user rating information to calculate the similarity that exists between objects or users, this can be either object-based or user-based. The most popular among the memory-based approaches is Collaborative Filtering (CF) [12]; this is a widely used framework that recommends services by relying on objects liked by similar users. Alejandro and Pablo [48] developed a CF approach based on subsequent matching by considering more information associated with the user. [49] used comments of users to implement an algorithm that dynamically suggests stories related to a given article. This approach has shown improvements in recommending services when a user comments content thread is considered. It is worthy of note that the performance of CF approaches, particularly the neighborhood approaches, heavily rely on the parameters used for recommending services. However, a drawback of using only the user-object rating matrix is that it offers limited information needed in predicting unknown ratings because of their inherent data sparsity. In an attempt to overcome this issue, Surya and Tripti [50] proposed a bi-clustering approach for neighborhood formation, by combining the object-based CF and user-based CF methods. Moreover, an implicit trust and distrust clustering-based CF was proposed in [51] which performs well under the sparsity issue, relatively larger datasets, and the cold start issues.

Furthermore, there are social-based recommender systems that depend on both content and social behavior of users to recommend items. They are categorized into matrix factorization-based methods and Probabilistic based methods. Due to their scalability and accuracy, Matrix Factorization (MF), a model-based method [52], has gained popularity since their success at the Netflix Prize competition. It has been widely used in the prediction of missing rating values in a user-object rating matrix because of its prediction efficiency and accuracy. To introduce certainty in the rating matrix prediction, Tadiparthi et al. [53] proposed a factorization framework that controls several inaccurate predictions by selecting suitable confidence levels depending on the application. It has been predominantly applied to the similarity matrix [54] to discover the similarities of users based on the different objects being rated by these users. The transitivity of social dependence in a social network is addressed using social-based matrix factorization models, as the dependency of a vector of a user's prominent attribute on a neighbor's attributes can be transmitted across the network, thereby making a given user attributes reliant on other users on the same network. To resolve the major problem of learning user's preference dynamics, Rafailidis et al. [55] proposed an MF framework that captures the dynamics of user preferences using a joint decomposition algorithm to extract the patterns temporarily created by users with multimodal user-object relations. Besides, Xu [1] resolves the problem of information overloading in social networks by proposing an MF technique that cluster users and consider other complex factors. Also, improving predictions and image retrieval through social tagging. Zechao et al. [56] proposed a system using a non-linear MF method that uses the priors of both inter-correlation as well as intra-correlations between tags and images to successfully make predictions on the relevance of tags to graphical content. Social based recommender systems have also been modeled using Bayesian networks. Yingjie et al. [23] developed a trust-based probabilistic approach to recommending products to users in the network using similarities among users computed based on attributes of products and inherent similarity among products. In [57], a multi-modal Bayesian embedding algorithm has been used to determine the influence of both sequential and temporal locations of persons on their check-in behaviors. Further, Devesh et al. [58] tackled the problem of chances of diffusion in communication signals alongside the social network links using a Bayesian network-based algorithm. Problem with regards to cold-start and data sparsity has been addressed to an extent in such recommendation methodologies. In studies by Juntao et al. [59], they modified the Bayesian Probabilistic matrix factorization model to incorporate social relationships and object contents fused through user ratings, which increased prediction accuracies.

Also, graph theory has been explored in modeling recommendation algorithms in social media. Yijun et al. [60] studied events recommendations using graph theory and proposed a novel event scoring algorithm by exploring the reverse random walk method to get a recommendation of events to users with higher precision and larger coverage. In [22], the cold-start problem of recommender systems has been resolved by a real-time algorithm that is centered on a graph embedding approach to generate user or object recommendations. Also, authors in [61] applied multi-modal-imaging, graph metrics based recommendations to measure the multi-layer network organization for cluster links, thereby allowing the assimilation of multiple networks to produce a single network. Influence measurement in social media has been a complex issue, [62] applied the bipartite graph by combining the user status and user-generated content in a network to measure the influence of diffusion. Chen et al. [63] analyzed the trustworthiness of information on mobile social networks and developed an algorithm using the social behavior of users, where relationships are rated and trust values between users are estimated based on their interaction evolution, level of contacts, and attributes of users.

There has been extensive use of diffusion algorithms too in developing effective recommendation systems using bipartite networks. However, a wide range of previous studies focused heavily on unidirectional diffusion (from rated objects to unrated objects), and undervalued the importance of the opposite direction. Guilin et al. [64] proposed a preferential mass diffusion method that is bidirectional by punishing the weights of popular objects. In the area of microblogging social networks, efficient services have been provided for information diffusion of news, ideas, and innovations. These services include the design of multi-level structures in analyzing the diffusion process of hot topics. Zhou et al. [65] designed a diffusion model that is based on a cascading framework using multilevel structures to retweeting network. Due to the uncertainty in human behavior, rumor diffusion on online social networks has caused irreparable damage to individual character and social cohesion. Some research works have addressed to minimize the damage caused to people on social media. In [25], a spatial-temporal diffusion framework has been designed by considering the effect of spatial diffusion, to limit the spreading of rumors within some pre-determined physical connections. Other researchers, including [66] looked at the social phenomena, such as viral diffusion of information between individuals, to determine the correlation among micro-level fundamental attributes of social networks and the emergence of such phenomena to avail effective, highly personalized recommendations to users according to their personalized needs and characteristics. Marketing diffusion systems have also benefitted from such strategies [67], to prevent information loss during the diffusion process. They developed a path planning method for advertisements that helps people with higher influence to spread information relating to marketing and also assist marketers to assess potential rewards under the various marketing policies. Jalali et al. [68] addressed the diffusion rate of online petition using system dynamics and postulate the essential processes of petition diffusion along with quantifying the same, and then predict the resultant dynamic model based on observed data. Their study leads to acknowledging the fact that users largely pull information towards themselves rather than it being a push to them. With the right people targeted, information can be easily diffused to the target users. Also, authors in [69] initially investigated the effects of the attributes of users as well as social networks in enhancing the diffusion process and put forward a better algorithm that improves the accuracy of predictions with the aid of neural network algorithms. However, An et al. [70] studied the recommender models from the perspectives of diversity and accuracy, and proposed an algorithm to enhance resource allocation similarities of diffusion-like models that improve diversity and accuracy simultaneously.

Despite the effectiveness of existing methods of service recommendation, there have been persistent problems affecting their accuracy and efficiency [71] with regards to user queries. Although some of the aforementioned algorithms improved recommendation accuracy, services offered by such models are not diverse enough to avail a range of options to users to choose from. A few other methods addressed to diversity issue but have not emphasized on accuracy. Therefore, a recommendation model that can effectively balance the trade-off between accuracy and diversity is still an essential requirement in OSN based recommender systems. In addressing this trade-off, this paper proposes a novel approach that resolves the accuracy-diversity dilemma in addition to resolving the data sparsity problem.

3. PROPOSED APPROACH

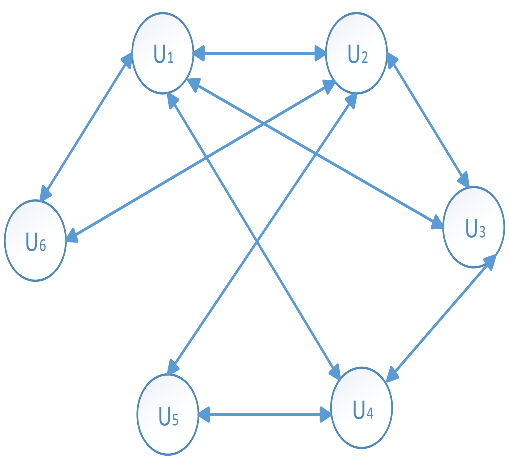
In this section, we first illustrate the social recommendation framework using a hypothetical example, then describe the friendship-based diffusion recommendation model that combines social information with item ratings to make recommendations. Finally, we describe the baseline methods.

3.1 The Social Recommendation framework

Consider the example illustrated in Figure 1(a) – (c). Figure 1(a) depicts a typical bidirectional friendship network graph with 6 users (nodes) and 9 relationships (edges) representing the interactions between the users. Each edge represents a connection between two users. Figure 1(b) is the friendship matrix formed using figure 1(a), where a value of 1 represents a connection between two users. In a coupled network, users would normally rate some objects (example, on a 5-point scale integer scale) to express their level of favor, as illustrated by the user-object bipartite graph in Figure 1(c), or be interested in objects tagged or rated by their friends. Such topics or objects are often discussed in the network according to the level of interests of users. Some topics or objects are normally overemphasized, and some may not be even noticed or rated. Now, the target of the proposed algorithm is to predict the ratings that a user might give to the unrated objects (missing values) using both the friendship matrix and the user-object matrix illustrated in Figure 1, and then recommend objects with higher ratings.

As illustrated in Figure 1(a), U_1 has a relationship with U_2, U_3, U_4, and U_6. Also, U_2 has a relationship with U_1, U_3, U_5, and U_6. U_2 rate objects O_2, O_5 and O_6 as shown in Figure 1(c). U_6 rates object O_3 aside O_2 and O_5. Now, the problem is that, if U_2 wants information about O_3 who would he/she ask or turn to. U_1 rates items O_1, O_3, and O_7 with ratings R_{11}, R_{13}, and R_{17} respectively. U_3 rates object O_4 aside O_1, O_3, and O_7. Although U_2 is a friend of U_1 and U_3, he/she would rather ask U_6 or U_5 since they rate some similar objects in common, thereby U_2 characterizes a higher similarity with U_6 and/or U_5 than that of U_1 and/or U_3. Again, U_2 would turn to U_5 for obtaining information about object O_7. The friendship matrix forms a nxn matrix as shown in Figure 1(b), with n number of users, while there are n users and m objects in the user-object network forming a $n \times m$ matrix, as shown in Figure 1(c).

1(a). A Typical Social Network



1(b). A Friendship matrix

	U_1	U_2	U_3	U_4	U_5	U_6
U_1		1	1	1		1
U_2	1		1			1
U_3	1	1		1		
U_4	1		1		1	
U_5		1		1		
U_6	1	1				

1(c). A user-Object Rating Matrix

	O_1	O_2	O_3	O_4	O_5	O_6	O_7
U_1	R_{11}		R_{13}				R_{17}
U_2		R_{22}			R_{25}	R_{26}	
U_3	R_{31}		R_{33}	R_{34}			R_{37}
U_4		R_{42}		R_{44}			
U_5	R_{51}		R_{53}		R_{55}	R_{56}	R_{57}
U_6		R_{62}	R_{63}		R_{65}		

Figure 1. An illustration of a social relationships and User-Object ratings in a coupled network

From the above illustration, it can be observed that each user would consult the most appropriate person among his/her friends' group, whilst making decisions on objects rather than randomly selecting users for help.

In our recommendation framework, we incorporate the scope of the relationship between users and then rank the candidate users based on their matching object interests and their social connections at top-K positions. To achieve this objective, we model the connections of users from these limited binary relevance data and propose a recommendation framework that includes two main phases. The first phase is the allocation of resources using a diffusion-based method, and the second phase includes ranking the objects that are yet to be noticed by users and recommending the top-K items to the target user.

3.2 The Friendship-based Diffusion Recommendation Method

Most diffusion-like recommendation methods work by allocating the initial level of resources to each object, α, in the user-object matrix. The resources assigned to these objects become a resource vector, which is then redistributed among those objects using various diffusion approaches. This redistribution continues until all the resources are distributed to the objects at the end of the diffusion process.

In our method, we explore the merits of the summation formula in Resource Allocation. First, for a given user u_x, $a_{u_x o_\alpha}$ denotes the initial resource that an object o_α is assigned. The value of $a_{u_x o_\alpha}$ is 1 if the user u_x collected object o_α and 0 otherwise, as shown in figure 2. In the second stage of the diffusion process, resources from the collected objects of user u_x are distributed to their neighbors using equation (1). We reduced the weights of less connected objects during the diffusion process, by modifying the summation formula using an inverse log function. In the modification process, we combined the inverse log function with the Resource Allocation index to produce the new summation formula, $h_y(x)$. This new formula is used to distribute the resources from object o_α to users that they are connected with.

$$h_y(x) = \sum_{o_\alpha=1}^{m} \frac{a_{u_y o_\alpha} a_{u_x o_\alpha}}{\log(k_{u_y} k_{o_\alpha})} \quad (1)$$

k_{o_α} is the number of users that select object o_α (i.e. the degree of the object o_α), k_{u_y} is the number of objects user u_y selects (i.e. the degree of user u_y) and *m* is the total number of objects. In a friendship network, user u_x and user u_y are friends if there exists an edge between node u_x and node u_y. We denote this as $A_{u_{xy}}$, where $A_{u_{xy}}$ =1 if an edge exists and 0 otherwise.

Now, the resources received by the various users in the network are redistributed back to their connected objects using both equations (2) and (3). However, in the second diffusion process as shown in the two equations, the effect of the objects and users with higher degrees are minimized. This is done to reduce the ranking of popular objects higher than not so popular objects.

$$h'_{o_\beta}(u_x) = \sum_{u_y=1}^{n} \sum_{o_\beta=1}^{m} \frac{a_{u_y o_\beta} h_y(y)}{k_{u_y} k_{o_\beta}} \quad (2)$$

$$h''_{o_\beta}(u_x) = \sum_{u_y=1}^{n} \sum_{o_\beta=1}^{m} \frac{a_{u_y o_\beta} h_y(y) A_{u_{xy}}}{K_{u_y} k_{o_\beta}} \quad (3)$$

K_{u_y} is the number of nodes that the user u_y is connected to in the social network (i.e. the social degree of user u_y). The final resources received by the items is a combination of the resources received through equations (2) and (3) using a tuning parameter μ as shown in equation (4). After receiving these resources, the items are ranked based on their final resources with the top-K objects that are not rated by the target user recommended.

$$h^* = \mu h'_{o_\beta}(u_x) + (1-\mu) h''_{o_\beta}(u_x) \quad (4)$$

The parameter μ varies the effect of users' social relationships as well as users' object degree in the diffusion process with an optimal value of μ given a better recommendation. Figure 2 illustrates an example of the diffusion process with different parameter values (μ=1, μ=0.5, μ=0).

3.3 Baseline Recommendation Methods

One of the earliest diffusion-based recommendation methods is the classic Mass Diffusion (MD) process in a user–object matrix proposed by Zhou et al. [72]. In their research, they put forward a Probabilistic Spreading algorithm using the resource transfer equation given by equation (5).

$$MD = \frac{1}{k_{o_\alpha}} \sum_{u_x=1}^{n} \frac{a_{u_x o_\alpha} a_{u_x o_\beta}}{k_{u_x}} \quad (5)$$

where $k_{o_\beta} = \sum_{u_x=1}^{n} a_{u_x o_\beta}$ and $k_{u_x} = \sum_{r=1}^{m} a_{u_x r}$ denote the degree of object O_β and user U_x respectively. The main advantage of MD is its higher rate of prediction accuracy; however, it characterizes a low diversity.

An additional diffusion-like model that mirrors a heat spre0ading algorithm is known as Heat Conduct (HC) [73]. This process uses the same initial resource vector as that of MD as shown in equation (6).

$$HC = \frac{1}{k_{o_\beta}} \sum_{u_x=1}^{n} \frac{a_{u_x o_\alpha} a_{u_x o_\beta}}{k_{u_x}} \quad (6)$$

The main variance between MD and HC is how resources are distributed in the network using the transfer matrix. MD employs an equal distribution of resources of each node to all the nearest neighbors of that corresponding node, with the overall resources remaining unchanged. However, in the case of HC, each node is presented with the same proportion of resources to that of their nearest neighbors; this increases the total amount of resources distributed in the process. HC characterize the advantage of higher diversity but low accuracy as compared to MD.

A few researches attempted to resolve this accuracy-diversity [74] problem with some effective algorithms. One of these algorithms is the hybrid model of HC and MD, named HHP [27], that is developed based on user-item relationship. HHP uses a tuning parameter λ to redistribute resources as shown in equation (7).

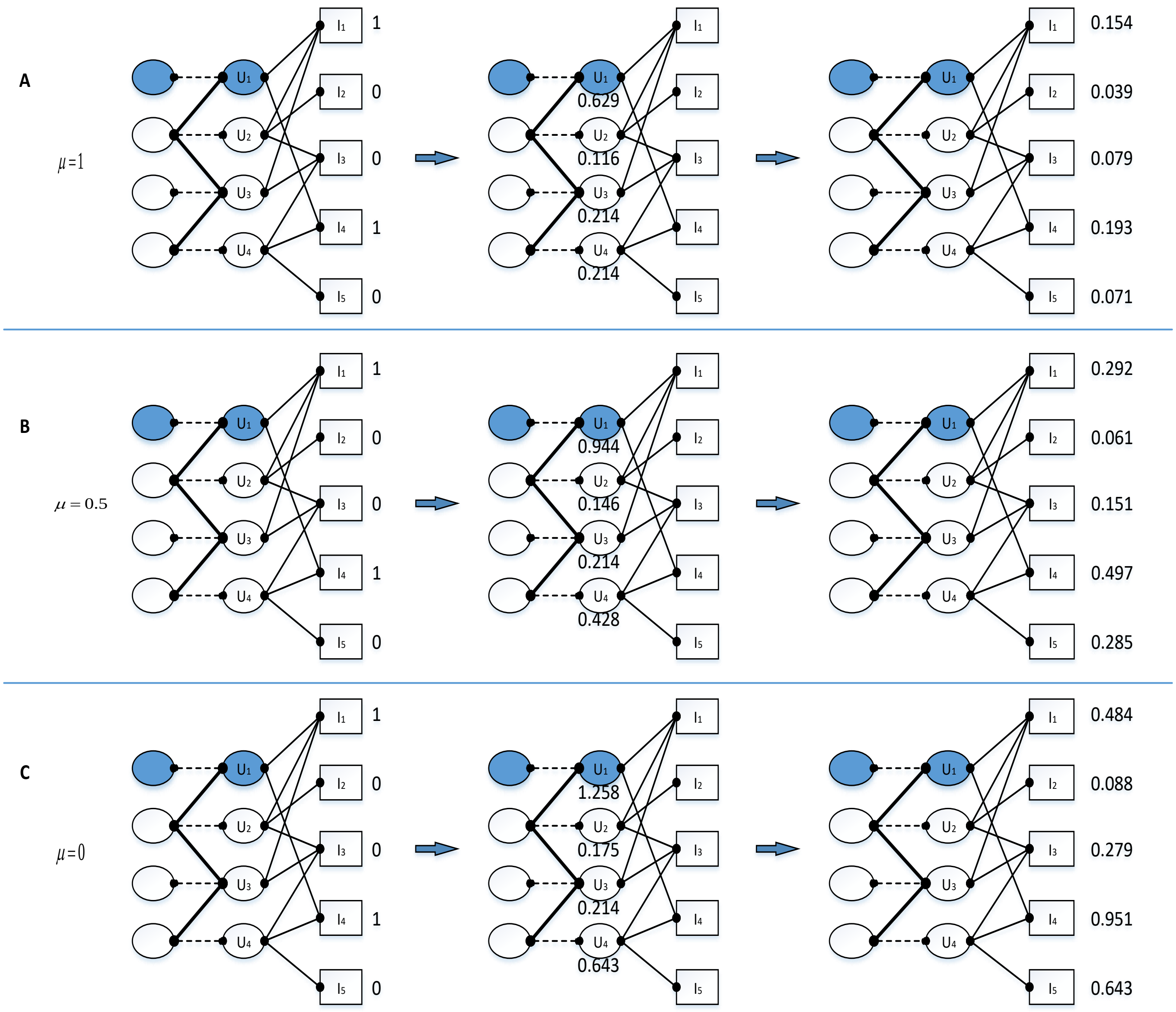


Figure 2. The FDR method with different parameter values. The shaded circle is the target user.

$$HHP = \frac{1}{k_{o_\alpha}^{1-\lambda} k_{o\beta}^{\lambda}} \sum_{u_x=1}^{n} \frac{a_{u_x o_\alpha} a_{u_x o_\beta}}{k_{u_x}} \tag{7}$$

In this algorithm, when λ = 1, the model works as a pure MD, and as a pure HC when λ = 0. So, by tuning λ between 0 and 1, a trade-off could be optimized with regards to accuracy and diversity.

However, in an attempt to incorporate social relationships into item recommendations, Xiaofang et. al [75] focused on general diffusion on coupled networks and proposed a diffusion-based algorithm, as shown in equation (8) by integrating the network of friends and user-item relations based on MD.

$$SocMD(o_\alpha, u_y) = p \sum_{u_x=1}^{n} \sum_{o_\beta=1}^{m} \frac{a_{u_x o_\alpha} a_{u_x o_\beta} a_{u_y o_\beta}}{k_{u_x} k_{o_\beta} k_{u_y}} + (1-p) \sum_{u_x=1}^{n} \sum_{u_l=1}^{n} \frac{a_{u_x o_\alpha} A_{u_{xl}} A_{u_{lj}}}{k_{u_x} K_{u_l} K_{u_y}} \tag{8}$$

where p is a resource ratio that is spread onto the user-object network, and the resources from a target user transferred to other users that are friends with him/her are given by the ratio (1-p). Each user transfers this ratio to their friends, which is then transferred finally to the objects in the user-object matrix.

To differentiate our method from SocMD shown in equation 8, our method combines the resource allocation formula with the inverse log function to punish popular items in the resource redistribution process. Also, our method diffusion process is one-way (i.e., from items to users, and users back to items), whiles the diffusion process in SocMD is two ways.

4. EXPERIMENT SET UP AND DATA EVALUATION

In this section, we preface our analysis with a description of; (1) the Experimental Datasets (section 4.1), (2) Evaluation Metrics (section 4.2), (3) Comparative baselines (section 4.3), (4) Experimental Settings (section 4.4), and (5) Results Analysis (section 4.5).

4.1 Experiment Datasets

We evaluate the performance of the FDR model against five other relevant baseline algorithms. The results are evaluated using two publically available real-world datasets. *Friendfeed* [46] and *Epinions* [45] obtained in [76].

Table 1: Datasets Summary

Datasets	Epinions	Friendfeed
Number of Users	4,066	4,148
Number of Items	7,649	5,700
User-Item ratings link	154,122	96,942
User-User link	217,071	386,804
User-Item Sparsity	0.00496	0.0041
User-User Sparsity	0.0131	0.0225

Friendfeed is a real-time social networking website that allows supplements for social bookmarking, social media, blogs and microblogging updates bookmarking, and other RSS/ATOM feeds, where people share customized feeds. The primary objective of this website is to assist users with tools needed in making decisions with regards to objects/information they are interested in. Epinions.com is a site where consumers review products, and was established in 1999. At Epinions, visitors could use reviews of objects to make decisions on their purchases. In both networks, users could follow each other based on their interests.

There are 4,066 users and 7,649 objects in the Epinions dataset, with 154,122 user ratings on objects and rating sparsity of 0.00496. Friendfeed, however, has 4,148 users, 5,700 objects, and 96,942 user ratings on objects with a rating sparsity of 0.0041. The two datasets also contain information about the social networks of users, which is essential in estimating the friendship relationships among users. Users establish friendships by rating objects and also by following one another in both the platforms. In creating the "user-user" friendship network, a link is formed between node U_x and U_y and assigned to 1 if user U_x is a friend to user U_y and 0 otherwise. The dataset in Epinions is made up of 217,071 friendship links with the user-user sparsity of 0.0131. Also, the dataset in Friendfeed is made up of 386,804 friendship links with a user-user sparsity of 0.0225 as illustrated in Table 1.

4.2 Comparative Baselines

To illustrate our models' performance improvement, we evaluate our model with four (4) other common diffusion methods: HC, MD, HHP, and SocMD. These methods are discussed further in Section 3.2 above.

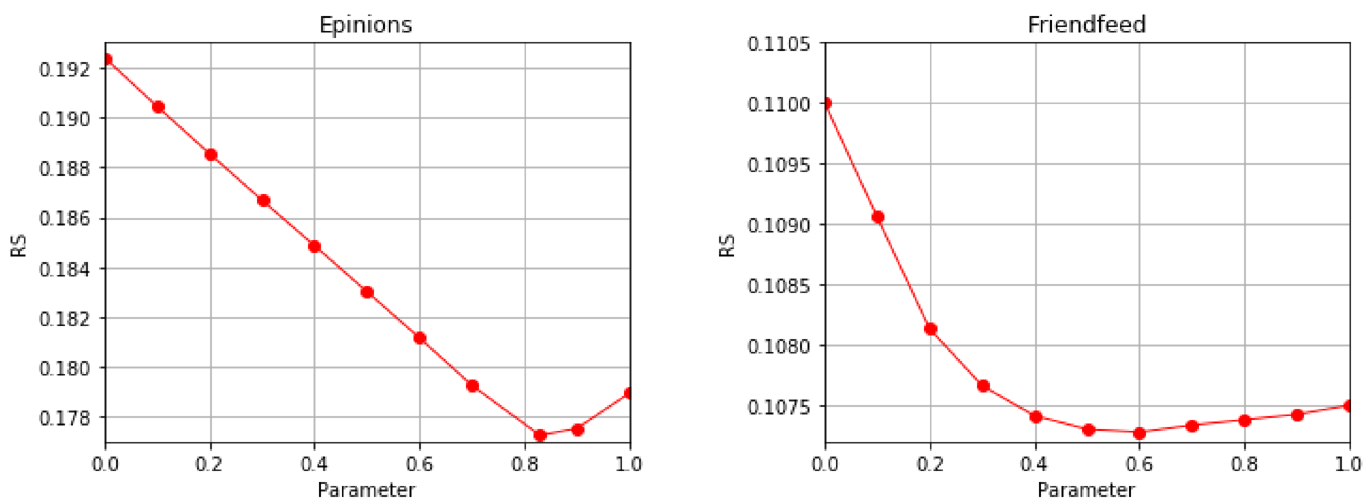
4.3 Evaluation Metrics

We evaluate our proposed FDR model performance based on five (5) accuracy metrics (i.e. Precision, Recall, F1, Ranking Score, and AUC), one diversity metric (i.e. Inter-Similarity) and one novelty metric (i.e. Average Degree) as detailed below.

Precision [64] measures the proportion of relevant objects that a target user selects to the total number of top-K objects recommended to the user. This is defined mathematically as;

$$P_{u_x}(K) = \frac{d_{u_x}(K)}{K} \tag{9}$$

where $d_{u_x}(K)$ denotes the total number of items collected by the target user U_x (i.e. relevant items of user U_x in the test dataset).



(a). Change in RS of the FDR method on Epinions

(b) Change in RS of the FDR method on Friendfeed

Figure 3. The Accuracy of FDR method as measured by the Ranking Score (RS) on both Friendfeed and Epinions datasets with scaling parameter from 0 to 1.

Recall [64] is the proportion of relevant objects selected to the total number of relevant objects. It estimates the probability that a relevant object is selected by a user. This is denoted mathematically as

$$R(K) = \frac{1}{m}\sum_{u_x=1}^{m}\frac{d_{u_x}(K)}{D(O_\alpha)} \tag{10}$$

where $D(O_\alpha)$ denote the number of all relevant items in the test set.

A higher precision value together with a higher recall value is an indication of higher accuracy in the recommendation, but these two measures often oppose each other as *P(K)* normally decreases with an increasing *K* while *R(K)* usually increases with increasing *K*. F1 metric is introduced in balancing the precision-recall trade-off.

The F1 Score [44] is used to balance the accuracy metric of both precision and recall, and is mathematically defined as:

$$F1 = \frac{2 * P(K) * R(K)}{P(K) + R(K)} \tag{11}$$

As a combination metric, the weights assigned to both recall and precision by the F1 score are even. A larger F1 value depicts a higher recommendation accuracy and vice versa.

Ranking score (*RS*) [72] is a metric used to evaluate the capability of recommendation methods to rank and recommend objects to a target user according to the users' liking. For every user-object link in the test set, the ranks of all objects in a user's recommendation list are generated. Taking an average of all the RS values of the individual user-object relationships in the test dataset to give an average ranking score that is used to predict the recommendation method's accuracy. The smaller the ranking score of an algorithm, the better the algorithms' performance and the higher the accuracy. Mathematically, given a target user U_x, the *RS* is denoted as

$$RS = \frac{1}{|E^P|} \sum_{(u_x, o_\alpha) \in E^P} \frac{p_{o_\alpha}}{k_{u_x}} \tag{12}$$

where $|E^P|$ represent the test set size, p_{o_α} denote the rank of object O_α in the recommendation list, and k_{u_x} denote the number of objects that are not collected/rated by the target user U_x in the training set. Smaller RS depict higher accuracy of a recommendation system.

The area under the receiver operating characteristic curve, (AUC), is a metric employed to also compute the accuracy of the recommendation model [44]. AUC represents the probability of a collected randomly chosen item characterizing a higher rank than an uncollected randomly chosen item in social network recommendation. After *N* number of independent comparisons, the resource received by both collected and uncollected item pairs is used to calculate the AUC value using equation (14).

$$AUC = \frac{1}{m} \sum_{u_x=1}^{m} \frac{(n_1 + 0.5 n_2)}{n}, \tag{13}$$

where *n* denotes the number of independent comparisons, n_1 denotes the total number of times for which the collected objects contain resources greater than the uncollected ones, and n_2 uncollected items characterize an equal number of resources for the target user U_x. A greater AUC value depicts a higher recommendation accuracy. If we generate scores from independent distribution, then the *AUC* will have a score of about 0.5, making it similar to a random recommendation. Therefore, a higher AUC value above 0.5 is an indication of how better the method is as compared to random choice.

The Inter-Similarity (IS) metric is used to measure the diversity of the recommendation method by computing the similarities between recommended objects [26]. The average IS for the recommender system is defined as

$$IS = \frac{1}{mK(K-1)} \sum_{\substack{o_\alpha, o_\beta \in E_p \\ o_\alpha \neq o_\beta}} CS(o_\alpha, o_\beta) \tag{14}$$

where $CS(o_\alpha, o_\beta)$ denote the similarity between objects o_α and o_β measured using the cosine similarity index, *K* is the length of the recommendation list of user U_x. A lower value of the IS metric is an indication of a higher diversity of the recommendation method.

Also the Average Degree (AD) [77] metric measures the popularity of objects in a user's recommendation list, and used to indicate the novelty of the recommender system. It is the summation of the degrees of all objects in the recommendation list to the number of objects in that list. This is defined mathematically as

$$AD = \frac{1}{mK} \sum_{o_\alpha=1}^{K} k_{o_\alpha} \tag{15}$$

A lower AD indicates a lower objects popularity and therefore a higher novelty of recommendation.

4.4 Experimental settings

In evaluating the performance of our proposed FDR model, we split each dataset into two sets randomly: a training set containing 90% of each dataset, and the remaining 10% forming the test (probe) set. The training dataset is made up of a user-object rating data and a user-user friendship network data. Most of the links (connections) in the friendship network are in the user-object network. We run MD, HC HHP, SocMD, and FDR methods on both the training sets of Epinions and Friendfeed datasets respectively. We then vary the parameters (i.e. λ, σ p, and μ) as defined in Section 3.2, from 0 to 1, with an increment of 0.1 in every iteration, to calculate the evaluation metrics outlined above. We repeat each experiment 10 times independently and obtain the average for an accurate result. In our proposed method, the parameter μ regulates the number of resources redistributed to the object network. A greater value of μ depicts greater resource distribution.

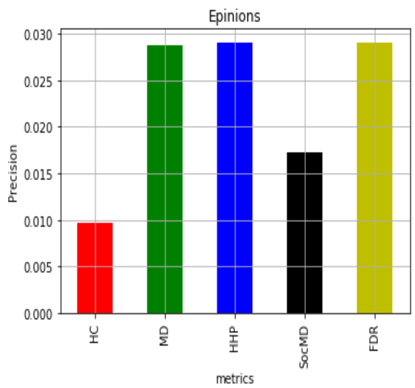
4.5 Results Analysis

In this section, we first illustrate the effect of the given parameter in our proposed FDR method, then analyze the performance of the FDR method against four other evaluation benchmarks.

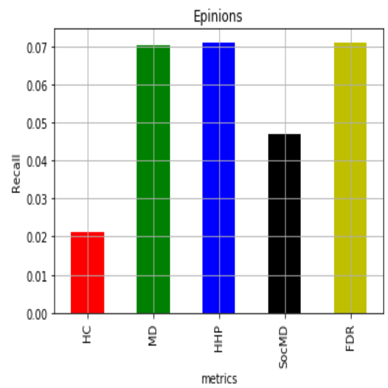
4.5.1. Effect of Scaling Parameter

By adjusting the tuning parameter μ from 0 to 1 we effectively capture the effects of user relations on resource distribution. This is evaluated on the Friendfeed and Epinions datasets using the Ranking Score (RS) evaluation metric as shown in Figure 3(a and b).

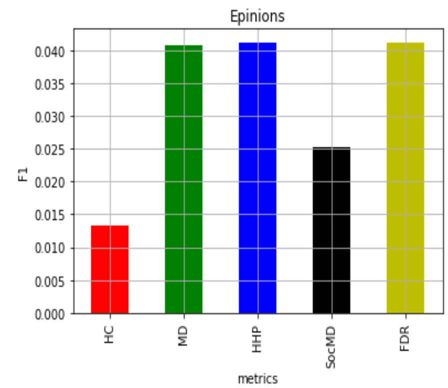
The FDR value reaches a lower RS score of ~0.177 when parameter μ is set between 0.8 and 0.9 as presented in Figure 3(a), on the Epinions dataset. Also, the FDR value reaches a



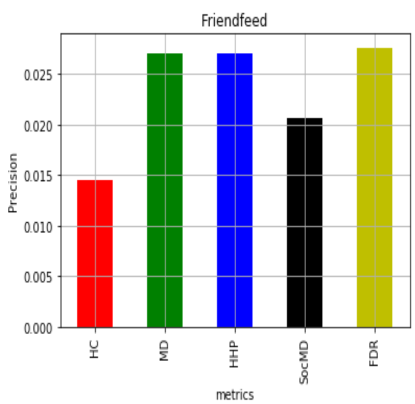
(a). Precision of the methods on Epinions dataset



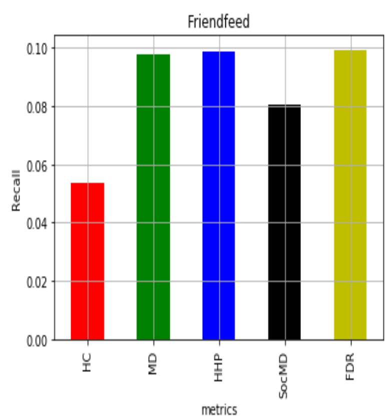
(b). Recall of the methods on Epinions dataset



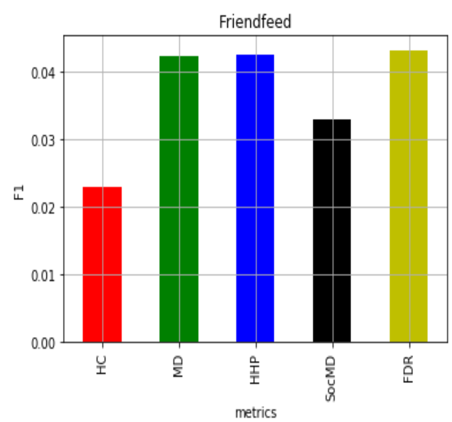
(c). F1 of the methods on Epinions dataset



(e). Precision of the methods on Friendfeed dataset



(f). Recall of the methods on Friendfeed dataset



(g). F1 of the methods on Friendfeed dataset

Figure 4. Comparison of the optimal values of Precision, Recall and F1 metrics of the various methods on both Epinions and Friendfeed datasets.

lower RS value of ~0.107 when parameter μ is set between 0.6 and 0.7 as shown in Figure 3(b) on the Friendfeed dataset. This depicts the universality of the optimal tuning parameter μ in both datasets.

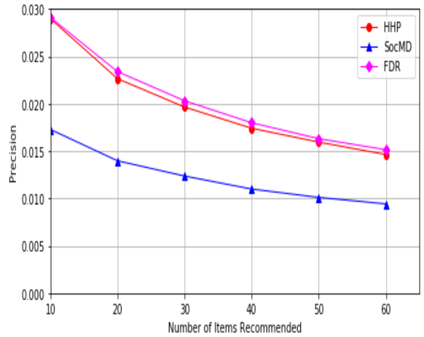
It can be further observed that the accuracy as measured by AUC increases as the parameter μ increases from 0 up until it reaches its maximum value, and then begins to decrease with any further increase in the value of μ in both the datasets. A sturdy decrease in RS value can be observed as parameter μ increases up until the RS value reaches its lowest possible value, and then begins to increase with a further decrease in parameter μ in both the datasets.

The effects of the scaling parameter μ suggest the importance of FDR in combining social relationships during object recommendations. With an optimal value of μ, the FDR algorithm can incorporate the higher accuracy characteristics of MD along with the greater diversity of HC to give us a better recommendation. The optimal value is obtained by averaging the recommendation length of results obtained under the FDR method for both Friendfeed and Epinions datasets.

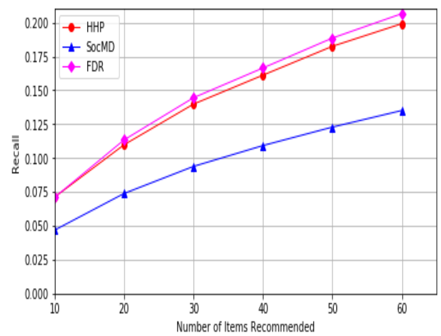
4.5.2. Recommendation Performance

This section provides a detailed evaluation of our proposed FDR algorithm against four existing models including HC, MD, HHP, and SocMD. The experimental results of recall, precision, and F1 scores for the proposed FDR and the other algorithms obtained under both the datasets with optimal recommendation parameters are illustrated in Figure 4(a-f). It can be seen that, except recall score in Epinions, all the other evaluated metrics of our proposed FDR model are greater than the scores of the baseline models. This implies that our recommender system can recommend objects for users with higher accuracy as compared with the other four methods.

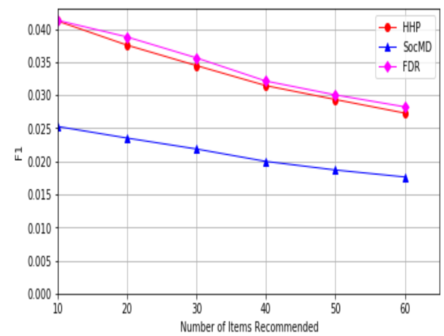
In particular, the precision, recall and F1 values of FDR measured against SocMD is higher by an average of 25%, 18% and 24% respectively on the Friendfeed dataset, and 40%, 34% and 38% respectively on the Epinions dataset as shown in figure 4(a)-4(f). SocMD performs poorly on sparse datasets, which leads to its declining performance against our proposed FDR method, as displayed in Table 1. Thus, our proposed FDR model exhibits better robustness of data sparsity, than the SocMD model. This means that the FDR method can be used to



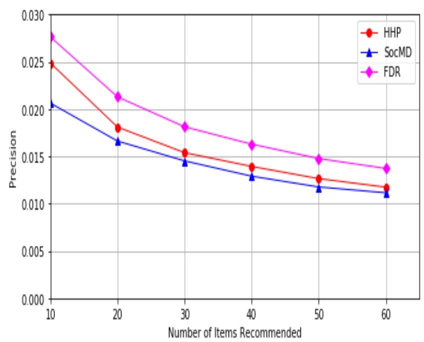
(a) Effect of Recommendation length on Precision (Epinions)



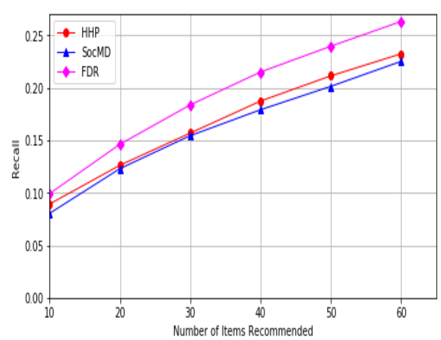
(b) Effect of Recommendation length on Recall (Epinions)



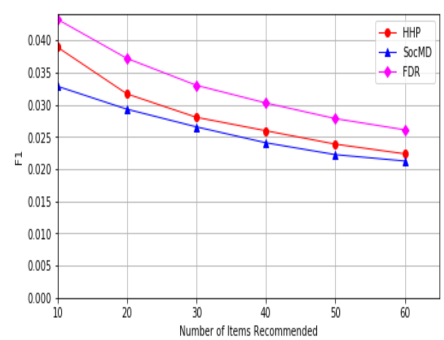
(c) Effect of Recommendation length on F1 (Epinions)



(d) Effect of Recommendation length on Precision (Friendfeed)



(e) Effect of Recommendation length on Recall (Friendfeed)



(f) Effect of Recommendation length on F1 (Friendfeed)

Figure 5. Measure of the effect of increase in recommended objects on the accuracy of recommendation

Table 2(a): Performance comparison of various diffusion models on Friendfeed dataset

	Friendfeed						
Method	AUC	RS	PR	RC	FI	IS	N
HC	0.873297	0.121494	0.014519	0.053501	0.022840	**0.026929**	**13**
MD	0.882639	0.105980	0.027014	0.097739	0.042329	0.063209	68
HHP	0.887982	0.104118	0.027100	0.098608	0.042516	0.045756	41
SocMD	0.887894	**0.100535**	0.020631	0.080492	0.032843	0.064163	81
FDR	**0.892954**	0.107277	**0.027637**	**0.099324**	**0.043242**	0.071664	38

Table 2(b): Performance Comparison of various diffusion models on Epinions dataset

	Epinions						
Method	AUC	RS	PR	RC	FI	IS	N
HC	0.776870	0.219914	0.009723	0.021329	0.013357	**0.018522**	**7**
MD	0.819668	0.177303	0.028781	0.070466	0.040870	0.070357	239
HHP	**0.826177**	0.178586	0.028987	**0.071190**	0.041193	0.069044	129
SocMD	0.787396	0.215883	0.017285	0.046827	0.025250	0.078725	260
FDR	0.816898	**0.177246**	**0.029114**	0.071026	**0.041299**	0.074833	171

solve the sparsity issues of recommender systems better than the SocMD model.

The performance evaluations of our proposed FDR model against all the four compared models are illustrated in Table 2(a and b). As shown in Table 2(a), our proposed FDR method achieves superior recommendations than all the other methods measured using the evaluation metrics, except the RS metric where the SocMD model slightly performed higher than the FDR model. The average RS values of our proposed FDR and SocMD are 0.107277 and 0.100535, respectively, for the

Friendfeed dataset. However, on the Epinions dataset in Table 2(b), the RS score of FDR is much better than all the other methods, with the RS value of SocMD method suffer a 17% decline to that of our proposed FDR method. FDR exhibits notable advantages to MD, HC, HHP and SocMD models as depicted by its larger AUC value of 0.892954 (shown in Table 2(a)) on the Friendfeed dataset and slightly lower with a score of 0.816898 (shown in Table 2(b)) on the Epinions dataset. However, the precision and F1 scores decrease with an increase in the number of recommended objects while the recall score increases with an increase in recommended objects in both datasets as shown by the three parameter-based methods in Figure 5.

Therefore, with regards to the recommendation accuracy, the proposed FDR model performs better than the other compared models based on experiments conducted using the Epinions dataset and also outperformed three of the other models on experiments performed using the Friendfeed dataset. Friendfeed is a denser network than Epinions as shown in Table 1. This suggests that our proposed FDR method can still exploit more precise social information in the Friendfeed dataset to exhibit better performance than the SocMD model. From the evaluation conduction on the sparse Epinions dataset, it can be established that the proposed FDR model exhibits greater robustness against data sparsity, in comparison with the other models.

With regards to the diversity of recommended objects, FDR can recommend objects with both higher and lower degrees, thereby improving the diversity of the recommendation as measured by the inter-similarity score and average degree score in both Friendfeed and Epinions datasets as shown in Table 2(a and b).

5. CONCLUSION

In this paper, we proposed a novel diffusion-based recommendation algorithm based on a friendship relationship that integrates social information into recommendation systems to increase the recommendation performance of recommendation systems for IoP. In our proposed method, we first modified the resource allocation formula by combining the resource allocation index with an inverse log function to reduce the weights of less connected objects during the first stage of diffusion. Then, the resources received by users are redistributed to the objects based on the friendship relations of the users. The resource is scaled by a tuning parameter before they are redistributed in the user-object network. We postulated an optimal scaling parameter that can be used to enhance the efficiencies of our proposed FDR method. Under this optimal range, FDR achieves reliable accuracy of recommendations in comparison with the compared four models of item recommendation. Experiments conducted on two different datasets and the corresponding outcomes validate the applicability of our proposed model for dynamic scenarios. The findings of this paper confirm the necessity of incorporating user relationships in a social network environment for an item recommendation, since such information provides useful insights for predicting user preferences to objects.

Our proposed FDR model enhances the recommendation accuracy and diversity by further looking at the users' social information asides the typically used user-item scores. Our model varies the fundamental distribution of object relationships, which can be of significant use in most of the recommender models for IoP without increasing the computational complexity. Experimental results, based on the Friendfeed and Epinions datasets, showed that, our proposed FDR method performs better than the benchmark methods, with a higher accuracy, better diversity, and encouraging novelty in recommendation. With regards to the effect of recommendation length on prediction accuracy, we discovered that, an increase in the recommendation length decreases the precision and F1 scores of our proposed method on both datasets, however, the recall values increase with an increase in recommendation length.

However, the present state of our proposed FDR method suffers a few limitations with regards to the datasets used and the modeling process in the experiments. This limitation calls for further investigations towards developing a better method that can effectively handle data density, sparsity, and accuracy. Although the evaluation of our model in this paper uses two rating datasets, they are only a representation of several online rating sites [78] and socio-economic systems. Thus, evaluating the performance of our proposed model in more rigorous and dynamic datasets is one of our future research plans. Moreover, our proposed method uses a simplistic approach of resource distribution, by scaling a tuning parameter in the diffusion process. Introducing the friendship relationships into rating-based approaches can potentially increase the recommendation performance of recommender systems, which is also a part of our future research plans. Also incorporating friendship relations into other features, and considering user profiling could be explored to improve personalization in recommender systems.

ACKNOWLEDGMENT

The work reported in this paper has been supported by National Natural Science Foundation of China Program (61502209 and 61502207), the Natural Science Foundation of Jiangsu Province under Grant BK20170069 and UK-Jiangsu 20-20 World Class University Initiative programme.

REFERENCES

[1] C. Xu, "A novel recommendation method based on social network using matrix factorization technique," *Information Processing & Management,* vol. 54, pp. 463-474, 2018/05/01/ 2018.

[2] F. A. Teixeira, F. M. Q. Pereira, H.-C. Wong, J. M. S. Nogueira, and L. B. Oliveira, "SIoT: Securing Internet of Things through distributed systems analysis," *Future Generation Computer Systems,* vol. 92, pp. 1172-1186, 2019/03/01/ 2019.

[3] Y. Kim and K. Shim, "TWILITE: A recommendation system for Twitter using a probabilistic model based on latent Dirichlet allocation," *Information Systems,* vol. 42, pp. 59-77, 2014/06/01/ 2014.

[4] J. Castro, J. Lu, G. Zhang, Y. Dong, and L. Martínez, "Opinion Dynamics-Based Group Recommender Systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* vol. 48, pp. 2394-2406, 2018.

[5] S. Deng, L. Huang, and G. Xu, "Social network-based service recommendation with trust enhancement," *Expert Systems with Applications,* vol. 41, pp. 8075-8084, 2014/12/15/ 2014.

[6] W. Wang, G. Zhang, and J. Lu, "Hierarchy Visualization for Group Recommender Systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* vol. 49, pp. 1152-1163, 2019.

[7] D. A. Adeniyi, Z. Wei, and Y. Yongquan, "Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method," *Applied Computing and Informatics,* vol. 12, pp. 90-108, 2016/01/01/ 2016.

[8] M. Gasparic and A. Janes, "What recommendation systems for software engineering recommend: A systematic literature review," *Journal of Systems and Software,* vol. 113, pp. 101-113, 2016/03/01/ 2016.

[9] M. Wang, L. Shi, L. Liu, M. Ahmed, and J. Panneerselvan, "Hybrid recommendation–based quality of service prediction for sensor services," *International Journal of Distributed Sensor Networks,* vol. 14, p. 1550147718774012, 2018.

[10] M. Ahmed, L. Liu, J. Hardy, B. Yuan, and N. Antonopoulos, "An efficient algorithm for partially matched services in internet of services," *Personal and Ubiquitous Computing,* vol. 20, pp. 283-293, June 01 2016.

[11] Y. Guo, L. Liu, Y. Wu, and J. Hardy, *Interest-Aware Content Discovery in Peer-to-Peer Social Networks* vol. 18, 2018.

[12] J. Aligon, E. Gallinucci, M. Golfarelli, P. Marcel, and S. Rizzi, "A collaborative filtering approach for recommending OLAP sessions," *Decision Support Systems,* vol. 69, pp. 20-30, 2015/01/01/ 2015.

[13] X. Feng, A. Sharma, J. Srivastava, S. Wu, and Z. Tang, "Social network regularized Sparse Linear Model for Top-N recommendation," *Engineering Applications of Artificial Intelligence,* vol. 51, pp. 5-15, 2016/05/01/ 2016.

[14] O. Kaššák, M. Kompan, and M. Bieliková, "Personalized hybrid recommendation for group of users: Top-N multimedia recommender," *Information Processing & Management,* vol. 52, pp. 459-477, 2016/05/01/ 2016.

[15] H. L. Z. Zhang, K. Liu, D. Wu, G. Zhang, J. Lu, "A hybrid fuzzy-based personalized recommender system for telecom products/services," *Inf. Sci.,* vol. 235, pp. 117-129, 2013.

[16] C. C. Y. Li, L. Lin, "A social recommender mechanism for location-based group commerce," *Inf. Sci.* , vol. 274, pp. 125-142, 2014.

[17] H. Zhang, W. Ni, X. Li, and Y. Yang, "Modeling the Heterogeneous Duration of User Interest in Time-Dependent Recommendation: A Hidden Semi-Markov Approach," *IEEE Transactions on Systems, Man, and Cybernetics: Systems,* vol. 48, pp. 177-194, 2018.

[18] A. T.-L. C. Porcel, M. Martínez, E. Herrera-Viedma, "A hybrid recommender system for the selective dissemination of research resources in a technology transfer office," *Inf. Sci. 184,* vol. 184, pp. 1-19, 2012.

[19] C. P. A. Tejeda-Lorente, E. Peis, R. Sanz, E. Herrera-Viedma, "A quality based recommender system to disseminate information in a university digital library,," *Inf. Sci.,* vol. 261, pp. 52-69, 2014.

[20] N. T. L. Son, "Intuitionistic fuzzy recommender systems: an effective tool for medical diagnosis," *Knowl. Based Syst.,* vol. 74, pp. 133-150, 2015.

[21] X. Zhu, H. Tian, G. Chen, and S. Cai, "Symmetrical and overloaded effect of diffusion in information filtering," *Physica A: Statistical Mechanics and its Applications,* vol. 483, pp. 9-15, 2017/10/01/ 2017.

[22] P. Liu, L. Zhang, and J. A. Gulla, "Real-time social recommendation based on graph embedding and temporal context," *International Journal of Human-Computer Studies,* 2018/03/01/ 2018.

[23] Y. Wang, G. Yin, Z. Cai, Y. Dong, and H. Dong, "A trust-based probabilistic recommendation model for social networks," *Journal of Network and Computer Applications,* vol. 55, pp. 59-67, 2015/09/01/ 2015.

[24] X. Wang, Y. Liu, G. Zhang, F. Xiong, and J. Lu, "Diffusion-based recommendation with trust relations on tripartite graphs," *Journal of Statistical Mechanics: Theory and Experiment,* vol. 2017, p. 083405, 2017/08/24 2017.

[25] L. Zhu and Y. Wang, "Rumor diffusion model with spatio-temporal diffusion and uncertainty of behavior decision in complex social networks," *Physica A: Statistical Mechanics and its Applications,* vol. 502, pp. 29-39, 2018/07/15/ 2018.

[26] T. Zhou, R.-Q. Su, R.-R. Liu, L.-L. Jiang, B.-H. Wang, and Y.-C. Zhang, "Accurate and diverse recommendations via eliminating redundant correlations," *New Journal of Physics,* vol. 11, p. 123008, 2009/12/04 2009.

[27] L. Lü and T. Zhou, "Link prediction in weighted networks: The role of weak ties," *EPL (Europhysics Letters),* vol. 89, p. 18001, 2010/01/01 2010.

[28] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, *et al.*, "Identification of influential spreaders in complex networks," *Nature Physics,* vol. 6, pp. 888-893, 2010/11/01 2010.

[29] X. Zhu, H. Tian, and S. Cai, "Personalized recommendation with corrected similarity," *Journal of Statistical Mechanics: Theory and Experiment,* vol. 2014, 05/16 2014.

[30] D.-C. Nie, Y.-H. An, Q. Dong, Y. Fu, and T. Zhou, "Information filtering via balanced diffusion on bipartite networks," *Physica A: Statistical Mechanics and its Applications,* vol. 421, pp. 44-53, 2015/03/01/ 2015.

[31] W. Ma, C. Ren, Y. Wu, S. Wang, and X. Feng, "Personalized recommendation via unbalance full-connectivity inference," *Physica A: Statistical Mechanics and its Applications,* vol. 483, pp. 273-279, 2017/10/01/ 2017.

[32] A. Fiasconaro, M. Tumminello, V. Nicosia, V. Latora, and R. N. Mantegna, "Hybrid recommendation methods in complex networks," *Physical Review E,* vol. 92, p. 012811, 07/14/ 2015.

[33] W. Zeng, Y.-X. Zhu, L. Lü, and T. Zhou, "Negative ratings play a positive role in information filtering," *Physica A: Statistical Mechanics and its Applications,* vol. 390, pp. 4486-4493, 2011/11/01/ 2011.

[34] L. Hu, L. Ren, and W. Lin, "A reconsideration of negative ratings for network-based recommendation," *Physica A: Statistical Mechanics and its Applications,* vol. 490, pp. 690-701, 2018/01/15/ 2018.

[35] E. Novak, Li, Q., "A survey of security and privacy in online social networks.," *Womencourage.acm.org,* vol. 9, pp. 1-12, 2012.

[36] L. Shi, L. Liu, Y. Wu, L. Jiang, and J. Hardy, "Event Detection and User Interest Discovering in Social Media Data Streams," *IEEE Access,* vol. 5, pp. 20953-20964, 2017.

[37] J. Su and T. C. Havens, "Quadratic Program-Based Modularity Maximization for Fuzzy Community Detection in Social Networks," *IEEE Transactions on Fuzzy Systems,* vol. 23, pp. 1356-1371, 2015.

[38] J. Yang, J. McAuley, and J. Leskovec, "Community Detection in Networks with Node Attributes," in *2013 IEEE 13th International Conference on Data Mining,* 2013, pp. 1151-1156.

[39] J. Li, X. Hu, L. Jian, and H. Liu, "Toward time-evolving feature selection on dynamic networks," in *Proceedings - 16th IEEE International Conference on Data Mining, ICDM 2016,* 2017, pp. 1003-1008.

[40] Y. Fang, R. Cheng, S. Luo, and J. Hu, "Effective community search for large attributed graphs," *Proc. VLDB Endow.,* vol. 9, pp. 1233-1244, 2016.

[41] X. Shen, H. Long, and C. Ma, "Incorporating trust relationships in collaborative filtering recommender system," in *2015 IEEE/ACIS 16th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD),* 2015, pp. 1-8.

[42] G. Guo, J. Zhang, and N. Yorke-Smith, "A Novel Recommendation Model Regularized with User Trust and Item Ratings," *IEEE Transactions on Knowledge and Data Engineering,* vol. 28, pp. 1607-1620, 2016.

[43] Y.-H. An, Q. Dong, C.-J. Sun, D.-C. Nie, and Y. Fu, "Diffusion-like recommendation with enhanced similarity of objects," *Physica A: Statistical Mechanics and its Applications,* vol. 461, pp. 708-715, 2016/11/01/ 2016.

[44] L.-J. Chen and J. Gao, "A trust-based recommendation method using network diffusion processes," *Physica A: Statistical Mechanics and its Applications,* vol. 506, pp. 679-691, 2018/09/15/ 2018.

[45] P. Massa and P. Avesani, "Trust-aware bootstrapping of recommender systems," in *Seventeenth European Conference on Artificial Intelligence*, Riva del Garda, Italy, 2006, pp. 29-33.

[46] F. Celli, F. Marta, L. Lascio, M. Magnani, B. Pacelli, and L. Rossi, "Social network data and practice: the case of friendfeed," in *Third International Conference on Social Computing, Behavioral Modeling, and Prediction*, Bethesda, MD, USA, 2010, pp. 346-353.

[47] X. H. J. Tang, H. Liu,, "Social recommendation: a review," *Soc. Netw. Anal. Min.,* vol. 3, pp. 1113-1133, 2013.

[48] A. Bellogín and P. Sánchez, "Collaborative filtering based on subsequence matching: A new approach," *Information Sciences,* vol. 418-419, pp. 432-446, 2017/12/01/ 2017.

[49] J. W. Q. Li, Y. Chen, Z. Lin, "User comments for news recommendation in forum-based social media," *Inf. Sci.* , vol. 180, pp. 4929-4939, 2010.

[50] S. Kant and T. Mahara, "Nearest biclusters collaborative filtering framework with fusion," *Journal of Computational Science,* vol. 25, pp. 204-212, 2018/03/01/ 2018.

[51] X. Ma, H. Lu, Z. Gan, and J. Zeng, "An explicit trust and distrust clustering based collaborative filtering recommendation approach," *Electronic Commerce Research and Applications,* vol. 25, pp. 29-39, 2017/09/01/ 2017.

[52] B. R. Koren Y., Volinsky C., "Matrix factorization techniques for recommender systems," *Computer,* vol. 42, pp. 30-37, 2009.

[53] T. V. R. Himabindu, V. Padmanabhan, and A. K. Pujari, "Conformal Matrix Factorization based Recommender System," *Information Sciences,* 2018/04/03/ 2018.

[54] P. Pirasteh, D. Hwang, and J. J. Jung, "Exploiting matrix factorization to asymmetric user similarities in recommendation systems," *Knowledge-Based Systems,* vol. 83, pp. 51-57, 2015/07/01/ 2015.

[55] D. Rafailidis, P. Kefalas, and Y. Manolopoulos, "Preference dynamics with multimodal user-item interactions in social media recommendation," *Expert Systems with Applications*, vol. 74, pp. 11-18, 2017/05/15/ 2017.
[56] Z. Li, J. Liu, and H. Lu, "Nonlinear matrix factorization with unified embedding for social tag relevance learning," *Neurocomputing*, vol. 105, pp. 38-44, 2013/04/01/ 2013.
[57] L. Huang, Y. Ma, Y. Liu, and A. K. Sangaiah, "Multi-modal Bayesian embedding for point-of-interest recommendation on location-based cyber-physical–social networks," *Future Generation Computer Systems*, 2017/12/27/ 2017.
[58] D. Varshney, S. Kumar, and V. Gupta, "Predicting information diffusion probabilities in social networks: A Bayesian networks based approach," *Knowledge-Based Systems*, vol. 133, pp. 66-76, 2017/10/01/ 2017.
[59] J. Liu, C. Wu, and W. Liu, "Bayesian Probabilistic Matrix Factorization with Social Relations and Item Contents for recommendation," *Decision Support Systems*, vol. 55, pp. 838-850, 2013/06/01/ 2013.
[60] Y. Mo, B. Li, B. Wang, L. T. Yang, and M. Xu, "Event recommendation in social networks based on reverse random walk and participant scale control," *Future Generation Computer Systems*, vol. 79, pp. 383-395, 2018/02/01/ 2018.
[61] K. Mandke, J. Meier, M. J. Brookes, R. D. O'Dea, P. Van Mieghem, C. J. Stam*, et al.*, "Comparing multilayer brain networks between groups: Introducing graph metrics and recommendations," *NeuroImage*, vol. 166, pp. 371-384, 2018/02/01/ 2018.
[62] Z. Zhu, J. Su, and L. Kong, "Measuring influence in online social network based on the user-content bipartite graph," *Computers in Human Behavior*, vol. 52, pp. 184-189, 2015/11/01/ 2015.
[63] S. Chen, G. Wang, and W. Jia, "Cluster-group based trusted computing for mobile social networks using implicit social behavioral graph," *Future Generation Computer Systems*, vol. 55, pp. 391-400, 2016/02/01/ 2016.
[64] G. Chen, T. Gao, X. Zhu, H. Tian, and Z. Yang, "Personalized recommendation based on preferential bidirectional mass diffusion," *Physica A: Statistical Mechanics and its Applications*, vol. 469, pp. 397-404, 2017/03/01/ 2017.
[65] Y. Zhou, B. Zhang, X. Sun, Q. Zheng, and T. Liu, "Analyzing and modeling dynamics of information diffusion in microblogging social network," *Journal of Network and Computer Applications*, vol. 86, pp. 92-102, 2017/05/15/ 2017.
[66] V. Arnaboldi, M. Conti, M. La Gala, A. Passarella, and F. Pezzoni, "Ego network structure in online social networks and its impact on information diffusion," *Computer Communications*, vol. 76, pp. 26-41, 2016/02/15/ 2016.
[67] Y.-M. Li, C.-Y. Lai, and L.-F. Lin, "A diffusion planning mechanism for social marketing," *Information & Management*, vol. 54, pp. 638-650, 2017/07/01/ 2017.
[68] M. S. Jalali, A. Ashouri, O. Herrera-Restrepo, and H. Zhang, "Information diffusion through social networks: The case of an online petition," *Expert Systems with Applications*, vol. 44, pp. 187-197, 2016/02/01/ 2016.
[69] L. Zheng-ren, L. Ting-jie, S. Wen-hua, and Z. Xiao-hang, "Predicting the scale of information diffusion in social network services," *The Journal of China Universities of Posts and Telecommunications*, vol. 20, pp. 100-104, 2013/08/01/ 2013.
[70] S. Zeng, "Link prediction based on local information considering preferential attachment," *Physica A: Statistical Mechanics and its Applications*, vol. 443, pp. 537-542, 2016/02/01/ 2016.
[71] Y. Wu, C. Yan, L. Liu, Z. Ding, and C. Jiang, "An Adaptive Multilevel Indexing Method for Disaster Service Discovery," *IEEE Transactions on Computers*, vol. 64, pp. 2447-2459, 2015.
[72] T. Zhou, J. Ren, M. Medo, and Y.-C. Zhang, "Bipartite network projection and personal recommendation," *Physical Review E*, vol. 76, p. 046115, 10/25/ 2007.
[73] Y.-C. Zhang, M. Blattner, and Y.-K. Yu, "Heat Conduction Process on Community Networks as a Recommendation Model," *Physical Review Letters*, vol. 99, p. 154301, 10/10/ 2007.
[74] L. Jiang, L. Shi, L. Liu, J. Yao, B. Yuan, and Y. Zheng, "An Efficient Evolutionary User Interest Community Discovery Model in Dynamic Social Networks for Internet of People," *IEEE Internet of Things Journal*, pp. 1-1, 2019.
[75] X. Deng, Y. Zhong, L. Lü, N. Xiong, and C. Yeung, "A general and effective diffusion-based recommendation scheme on coupled social networks," *Information Sciences*, vol. 417, pp. 420-434, 2017/11/01/ 2017.
[76] D. Nie, Z. Zhang, J. Zhou, Y. Fu, and K. Zhang, "Information filtering on coupled social networks," *PLoS ONE*, vol. 9, p. e101675, 2014.
[77] Z.-K. Zhang, C. Liu, Y.-C. Zhang, and T. Zhou, "Solving the cold-start problem in recommender systems with social tags," *EPL (Europhysics Letters)*, vol. 92, p. 28002, 2010/10/01 2010.
[78] J. Gao and T. Zhou, "Big data reveal the status of economic development," *Dianzi Keji Daxue Xuebao/Journal of the University of Electronic Science and Technology of China*, vol. 45, pp. 625-633, 07/30 2016.

Diyawu Mumin received his B.Sc. degree from the University of Ghana, Accra, Ghana, in 2006 and an M.Sc. degree from Sikkim Manipal University, Manipal, India, in 2011. He is currently working towards the Ph.D. degree at the School of Computer Science and Telecommunication Engineering, Jiangsu University, Zhenjiang, China. His research interests include Service Computing, Link predictions, Data mining, and Online Social Networks.

.

Lei-lei Shi received the B.S. degree from Nantong University, Nantong, China, in 2012, and the M.S. degree from Jiangsu University, Zhenjiang, China, in 2015. He is currently working towards the Ph.D. degree at the School of Computer Science and Telecommunication Engineering, Jiangsu University, Zhenjiang, China. His research interests include Event Detection, Data Mining, Social Computing, and Cloud Computing.

Lu Liu is the Head of the School of Informatics and Professor of Informatics at the University of Leicester, UK. Prof. Liu received his Ph.D. degree from the University of Surrey and M.S. degree from Brunel University. Prof. Liu's research interests are in the areas of data analytics, service computing, cloud computing, Artificial Intelligence, and the Internet of Things. He is a Fellow of British Computer Society (BCS).

John Panneerselvam is a Lecturer in Computing at the University of Leicester, United Kingdom. John received his Ph.D. in Computing from the University of Derby in 2018 and an MSc in advanced computer networks in 2013. He is an active member of IEEE and British Computer Society, and a HEA fellow. His research interests include cloud computing, fog computing, Internet of Things, big data analytics, opportunistic networking and P2P computing. He has won the best paper award in IEEE International Conference on Data Science and Systems, Exeter, 2018.