IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 15 February 2023, accepted 16 March 2023, date of publication 28 March 2023, date of current version 19 April 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3262587

RESEARCH ARTICLE

Co-Evolving Popularity Prediction in Temporal Bipartite Networks: A Heuristics Based Model

KHUSHNOOD ABBAS[1,2], MOHAMMAD KAMRUL HASAN[3], (Senior Member, IEEE), ALIREZA ABBASI[2], SHI DONG[1], TAHER M. GHAZAL[3,4,5], SITI NORUL HUDA SHEIKH ABDULLAH[3], ASIF KHAN[6], (Member, IEEE), DABIAH ALBOANEEN[7], FATIMA RAYAN AWAD AHMED[8], THOWIBA E. AHMED[7], AND SHAYLA ISLAM[9], (Member, IEEE)

[1]School of Computer Science and Technology, Zhoukou Normal University, Zhoukou 466001, China
[2]School of Engineering and IT, The University of New South Wales (UNSW), Sydney, NSW 2052, Australia
[3]Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor 43600, Malaysia
[4]Applied Science Research Center, Applied Science Private University, Amman 11937, Jordan
[5]School of Information Technology, Skyline University College, University City Sharjah, Sharjah, United Arab Emirates
[6]Department of Computer Application, Integral University, Lucknow, Uttar Pradesh 226026, India
[7]Computer Science Department, College of Science and Humanities, Imam Abdulrahman Bin Faisal University, Jubail 31961, Saudi Arabia
[8]Computer Science Department, College of Computer Engineering and Science, Prince Sattam Bin Abdulaziz University, Al-Kharj 16273, Saudi Arabia
[9]Institute of Computer Science and Digital Innovation, UCSI University Malaysia, Kuala Lumpur 56000, Malaysia

Corresponding authors: Mohammad Kamrul Hasan (hasankamrul@ieee.org), Siti Norul Huda Sheikh Abdullah (snhsabdullah@ukm.edu.my), and Asif Khan (asif05amu@gmail.com)

This work was supported by Universiti Kebangsaan Malaysia under Grant LRGS/1/2019/UKM-UKM/2/7.

ABSTRACT One of the big challenges of our modern life is to find the right items or contents on the Internet and particularly in social media. One way of addressing the information overload problem in social media is to predict the future trends and popularity of online items. The popularity of an item can be measured by its attractiveness, i.e., the number of times it is being used. This popularity prediction can be translated to a link prediction and ranking problem, which aims to predict the link gain of the items in a user-item interaction network. User-item interactions in an online environment can be modelled as a bipartite network, where a link represents an event, reflecting a user buys or collects an item. Popularity prediction problem in temporal bipartite networks is of great interest to researchers. In this study, we propose a heuristic based model which only consider nodes collective link gain in a recent past time window of time as well as total link gain. To evaluate our model's efficiency, we tested them on co-evolving social media items. We also evaluated the models' performance on five information retrieval metrics (i.e., Area Under the Receiver Operating Characteristic, Kendall's rank correlation tau, Precision, Novelty, and temporal novelty). The proposed model does not need hyper-parameter learning, which makes it the best choice for highly temporal and data streaming scenarios.

INDEX TERMS Temporal bipartite networks, ranking, popularity prediction.

I. INTRODUCTION

Recent years have seen the quick advancement of online services such as social networking, news, blogging, and so on. The fast development of information and communication technologies, and in particular, the ease of access to the Internet through portable devices such as smartphones, creates a sheer volume of data about users, items, and different types of interactions among them like buying or viewing items by users [1], [2]. Such enormous data conveys 'information overload' problem, under which it is hard to find, appropriate results or choices. This requires numerical ways to deal with ranking the accessible choices; otherwise stated as predicting the most appropriate items to users' need or question [3]. This problem can be converted to popularity or trend prediction, especially in internet-based

The associate editor coordinating the review of this manuscript and approving it for publication was Claudio Zunino.

services. One method for solving this issue is by applying a network science approach [4], [5], [6], [7], [8], [9], [10] by converting user item interaction to a bipartite network which applied in recommendation systems [11]. User item or item-item interaction can also be solved as a link prediction problem in evolving networks [12], [13]. Though we will focus in this study only on ranking nodes in dynamic networks.

The ability to predict an item's (or user's) popularity has a wide range of applications, including studying people's social influence or using it for advertising purposes. Popularity can be determined by how desirable an item is or how frequently it is used. In real-world situations, it is important to also consider connections between entities of different types in addition to connections between entities of the same type. As such, the network of user-item interactions can be depicted as a bipartite network that evolves over time, where users and items are two distinct types of nodes and the connection between them represents an interaction, such as when a user likes or buys an item. For instance, on social media platforms like Instagram [14], Digg, YouTube, and Movielense, viewing activities create a bipartite connection where the edges signify user's reading or viewing activities of news, videos, and movies. In temporal bipartite networks, the popularity prediction problem can be transformed into a future link prediction problem by using graph data structures. Identifying the rise and fall of an item's popularity dynamics allows for the examination of temporary user-item interactions. It can be challenging to infer general user interests from an item's characteristics because consumers often do not explicitly state an item's content level attributes in an online context (e.g. video upload on YouTube). Therefore we need to rely on other features which can be easily obtained. Many real-world issues, including stock market dynamics [15], have been successfully solved using longitudinal social media data, for examples, see [16], [17], [18], [19], and [20]. Short-term trends and popularity also offer some insight into general online user behavior. As a result, these models can be used by various research communities. It is important to consider macro-level analysis when predicting whether items are gaining or losing popularity over time. Although there has been a lot of study on predicting the popularity of specific items, it has not been successful in identifying newly popular items at the system level.

As was previously discussed, ranking nodes according to their importance or potential popularity is a fundamental challenge. User-item interactions can be described as temporal bipartite networks. Studies that have already been done on node ranking in networks have primarily focused on mono-partite and static networks, ignoring temporal factors [4]. It is apparent that when networks change over time, node importance will change as well because many systems expand or change over time. Only in an ideal rich-gets-richer setting where the oldest node is the most popular can static metrics function. But in reality, systems evolve under many constraints such as aging, deletion of nodes, and so on. Therefore, time-dependent algorithms are proposed to consider the evolutionary aspect of networks. These algorithms primarily aim to reduce the dominance of older nodes, which helps in giving newer nodes a higher ranking. Various mathematical operations can be used to analyse the ageing process in evolutionary systems, including exponential decay [10], [21], [22], [23], power-law decay [24], [25], [26], and log-normal decay [27], [28].

The remainder of this paper is organized as follows. Section [II] reviews the challenges of predicting popular items and ranking nodes as an approach to address this issue. Our proposed methodology and algorithms for predicting nodes' popularity and two existing models for popularity prediction (i.e. preferential attachment-based and temporal-based models) are presented in the section [III]. Section[IV] presents results of applying our proposed methods to two real datasets, i.e. Movielens and Diggs. Finally the paper concludes summarizing our findings, discussing research limitations and future works.

II. RELATED WORK

Bakshy et al. and Martin et al [29], [30] reported it is difficult to predict which item will be more popular than others even having every minute details of items and features about the person who is sharing it. The issue turns out to be increasingly tractable when we are permitted to look into the early history of the items' popularity. The most important finding is that early history information about the speed of selection, attributes of individuals who interacted with it and the associations between them can anticipate the items' popularity. This intuition shows positive results for both anticipating the eventual popularity of an item [24], [31] and whether the item will reach the top of 50% of popular items or not [32], [33]. In their study, researchers Bentley et al. [34], [35] proposed a model that examined the impact of users copying recent behavior of their peers. They found that their model was able to effectively generate real data, and that items that have recently gained popularity are likely to continue to do so in the future. This result deviates slightly from the well-known "rich-gets-richer" effect, which can be attributed to the aging effect in social and cultural artifacts [9], [10], [36]. According to a recent study by Candia et al. [36], the social and cultural values of products decay bi-exponentially. Other studies have also found that some items gain popularity multiple times [37]. These studies have identified various factors that affect the recurrence of cascades, including the structure of the network, initial burst size, temporal patterns, sharer's characteristics, and multiple copies in the network. Among these factors, temporal patterns, particularly initial bursts that exceed a certain threshold, play a crucial role. Additionally, novelty [38] is another key driving factor for popularity. Novelty can lead to the dissemination of both false and true information on social media [39]. While predicting which items will have future impact at their creation time is a challenging task, having some historical data on the items allows for early predictions of future popularity [5], [9], [40]. Liebig et al. [41] presents a new approach to predicting the future popularity of new items in rating, (user-item) networks, using a bipartite clustering coefficient. The method is applied to predict the popularity of movies in the MovieLens network and stories in the Digg network, with a success rate of over 65% for the former and over 50% for the latter, outperforming existing methods. Lin et al. [42] proposed Dynamic Activeness (DA) model is based on the concept of activeness and captures the three characteristics of a trend: intensity, coverage, and duration. The method uses a combination of network features and historical user activity data to predict the future popularity of topics in a social network. Gao et al. [43] presents a method for predicting the popularity of posts in a microblogging network. The authors propose a method that uses easily obtainable features, such as the number of followers and the frequency of posts, to predict the popularity of posts. They have solved it as binary classification problem on the basis of some threshold value. Mishra [44] presents a method for predicting the popularity of posts on social media. They propose a new model that bridges the gap between feature-driven and generative models by modeling social cascades with a marked Hawkes self-exciting point process and a predictive layer on top. The authors also propose a recurrent neural network-based model for asynchronous streams to account for multiple sources of external influence. This work provides accurate and explainable popularity predictions.

A. POPULARITY AS A NETWORK GROWTH MODEL

As user item interactions can be modeled as a bipartite network therefore to exploit network structural features many researchers solved this problem as a network growth model also [5], [9], [40], [45], [46]. The paper [45] introduces a new network growth model that addresses the disadvantage faced by late arrivals in the preferential attachment model. The model incorporates a recent-degree-change bias to give an attachment probability boost to nodes with higher relative degree change, characterizing a hot-get-richer mechanism. The proposed model produces later high-ranking nodes compared to the PA model and, under certain parameters, produces network structures similar to PA networks. Further Pongnumkul et al. [46] studies the behavior of music consumers on online music streaming services through an evolution model and analysis. The paper proposes a new bipartite growth model to understand the evolution of online music services by adding fitness and aging functions to the existing bipartite network model. Chandra et al. [47] proposes a growth model for online emerging user-object bipartite networks to study the selection behavior of web users. The network evolves with the constant growth of both users and objects and the arrival of edges from the user set, with a combination of preferential and random attachment mechanisms for external edges from new users and internal edges from old users. Zhou et al. [48] proposed an age-based diffusion model that takes into account the impact of time on the diffusion of information about a publication. The method considers both the aging of the publication and the temporal dynamics of the citation network. Researchers in [5], [9], and [40] has solved popularity prediction problem as ranking in evolving bipartite network model. The have considered many user-item interaction scenarios such as user-movie, user-news item, and user-Facebook wall-post interactions etc. and gave solutions based on node ranking.

B. IF INFLUENTIAL NODE BECOME THE POPULAR FUTURE NODES

Finding influential nodes by ranking in static networks is a well-known problem researched by the physics community. Here we are discussing whether influential nodes become the most popular nodes in a dynamic network scenario. Although our work focuses on evaluating algorithms based on popularity. The in-degree (degree in un-directed networks) is the most basic measure of a node's popularity or importance prediction. Ranking nodes according to their degree appears to be an effective strategy for large-scale networks and has a low computing cost, but it has poor accuracy. Other centrality measures, such as betweenness centrality and closeness, have been suggested to address the inadequacies of the degree centrality. The concept of breaking down the network has been pioneered using techniques like K-Shell [49]. Following the concept of decomposition, several methods have used K-Shell to rank nodes according to importance for example, gravity centrality [50] and extended K-Shell sum [51]. Degree centrality, local clustering coefficient, and k-shell decomposition are recently coupled [52] to capture both local and global importance. Kitsak et al. [49] demonstrated in a study that nodes in the core region of a network are more likely to be influential. This work served as the basis for Chen et al. [53] coreness centrality measure. The coreness approach uses K-Shell decomposition to assign numerous nodes to the same shell. However, nodes in the same shell may have varying degrees of influence. Additionally, in core-less networks, K-Shell may be unable to recognise influential nodes.

C. CONSIDERING TEMPORAL EFFECT IN RANKING

To consider temporal biases in rankings based on in-degree, Newman used a rescaled version of in-degree for paper citation networks. The rescaled version normalizes citation (in-degree) for papers of similar age. Researchers in [54], and [55] have proposed re-scaled equation by average citation counts of papers in the same area and in the same year. Likewise, PageRank (a well-known ranking algorithm) is also found to have a bias towards selecting old nodes as more important [56]. Some researchers proposed a re-scaled PageRank model that considers a moving temporal window Δt. Δt is the only parameter to re-scale and can also be considered as tuning the ranking between old and new nodes. When Δt is large the model will rank old nodes higher and when Δt is smaller it results in biased ranking towards temporally fluctuating item which can be noise also. In effect a node could be ranked higher while it has very low degree. Both results suggest that consideration of Δt should be approached with caution to avoid both extremes. Due to their great computational complexity, not all of these approaches are scalable to large-scale networks. Further Zeng et al. [5] proposed Popularity-Based Predictor which is convex combination of node current degree (considering user-item bipartite network) and recent degree increase. These findings give us hint that a simple re-scaled equation can be a good predictor for future popularity of items. Therefore in this work we propose to consider in-degree as well as in-degree up to recent past time window concept to predict future popular as well as novel popular nodes.

One research gap in the field of dynamic graph analysis is the lack of effective models that can accurately predict the future popularity of nodes in a dynamic network while considering the temporal dynamics of the network. Current methods for dynamic node ranking often rely on static features of the network, such as degree centrality or PageRank, which may not capture the temporal changes in the network and the underlying patterns that drive these changes. Additionally, many existing models do not take into account the historical popularity of nodes, which can be an important factor in predicting future popularity. Furthermore, most of the current models for dynamic node ranking have high time and computational complexity which makes them impractical for large-scale dynamic networks, especially in real-time applications. Therefore, there is a need for new models that can effectively capture the temporal dynamics of the network and accurately predict the future popularity of nodes while taking into account historical popularity and having low computational complexity.

III. METHODS

This research aims to predict the popularity of objects (items) for a user based on the previous and existing interest users have shown on objects, modeling their interaction via a bipartite network including users ($u \in U$) and objects ($o \in O$) as two separate groups of nodes and links among any pair of '(user, object)' representing the use of an object by a user. This network can be represented in the form of an adjacency matrix (A_{uo}), in which each cell shows the existence of a link between an object and a user (i.e., the use of o by u). The degree of an object (k_o) the number of users who have used or collected that object (o), has been widely considered as a proxy for its popularity or attractiveness. As we are aiming to predict the future popularity of the objects, we use the time-based adjacency matrix, $A_{uo}(t)$ to calculate an object's degree at a previous or current time to predict its future degree. The degree of object o at time t is given by:

$$k_o(t) = \sum_u A_{uo}(t) \tag{1}$$

$$\Delta k_o(t, t+T_F)) = k_o(t+T_F) - k_o(t) \tag{2}$$

where $k_o(t)$ is the link total link gain by object o upto time t. Similarly, $\Delta k_o(t, t+t_F)$ is the link gain of an object during future time window T_F.

Before describing our proposed model for popularity prediction, we will briefly discuss two basic benchmark models which serve as the basis of our approach and are also used for comparing the performance of our model with.

A. THE PREFERENTIAL ATTACHMENT-BASED MODEL

The preferential attachment model, also known as Barabasi-Albert (BA or PA) model [57], is a very simple generative model that posits that a node's current degree is a good predictor of future link gain. In other words, the PA model predicts the future popularity of object o as the fraction of its current degree. Thus, the future link gain probability of object o (or prediction score for object o at any time t) based on PA model is given by $s_o(t)$-

$$s_o(t) \sim \frac{k_o(t)}{\sum_{u,\forall o} k_o(t)} \tag{3}$$

where, $k_o(t)$ reflect the degree of object o at the time (t), and the denominator is used to calculate the sum of the degrees of all objects in the network.

B. TEMPORAL-BASED PREDICTOR (TBP)

The temporal-based model, proposed by Zhou et al. [40], considers link influence decay over time effect when calculating an object's probability score for future link gain.

$$s_o(t) = \sum_u A_{uo}(t) \exp(\gamma(T_{uo} - t)) \tag{4}$$

where $s_o(t)$ is prediction score for object o at time t, $A_{uo(t)}$ is the user-object adjacency matrix at time t, and T_{uo} is the time when the link between o and u is formed, $T_{uo} < t$, and γ is the link influence decay rate.

C. OUR PROPOSED MODEL: NON-PARAM

We are proposing a new model for future popularity prediction of an object. As presented in Fig. 1, we must consider three time windows: 1) past time window (time before $t - T_P$); 2) the recent time window (T_P), up to present time t; and 3) the immediate future time window (T_F). Since some items' popularity increases and fade many times [37], we should consider both past and recent (current) time periods. Therefore, we use the following three-rank

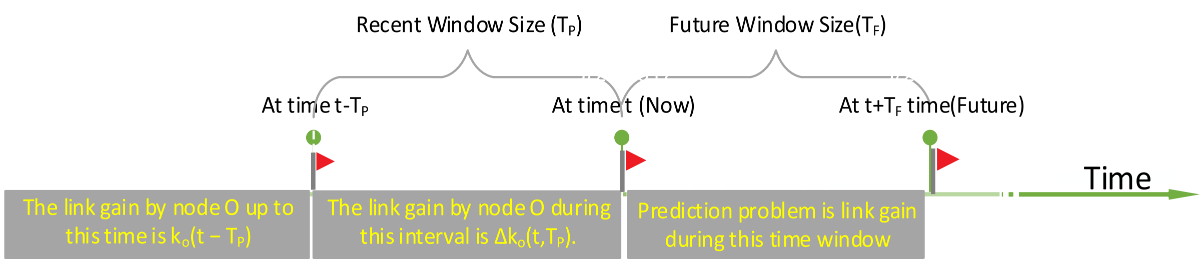

FIGURE 1. **The above figure describes the prediction problem for a ranking in evolving networks. The time is increasing from left to right. The current time is *t*. This figure shows how the time window concept is used in our proposed work.**

scores for each time windows:

$$K^P_{so}(t, T_P) = \frac{k_o(t - T_P)}{\sum_{u,\forall o} k_o(t - T_P)} \tag{5}$$

$$K^C_{so}(t) = \frac{k_o(t)}{\sum_{u,\forall o} k_o(t)} \tag{6}$$

$$K^R_{so}(t, T_P) = \frac{\Delta k_o(t, T_P)}{\sum_{u,\forall o} \Delta k_o(t, T_P)} \tag{7}$$

where $K^P_{so}(t, T_P)$, $K^C_{so}(t, T_P)$ and $K^R_{so}(t, T_P)$ are the rank score based on objects' link gain up to the past time window, up to current time t, and during the recent time window $(t - T_P,t)$ respectively. $\Delta k_o(t, T_P) = k_o(t) - k_o(t - T_p)$, $k_o(t)$ reflect the link gain up to time (t) Fig. 1. In order to model our problem we have $R^P_{so}(t, T_P)$, $R^C_{so}(t, T_P)$ and $R^R_s(t, T_P)$ rank scores vectors of nodes as follows-

$$R^P_{so}(t, T_P) = \sum_{K^P_{so'}(t,T_P) <= K^P_{so}(t,T_P)} K^P_{so'}(t, T_P) \tag{8}$$

$$R^C_{so}(t) = \sum_{K^C_{so'}(t) <= K^C_{so}(t)} K^C_{so'}(t) \tag{9}$$

$$R^R_{so}(t, T_P) = \sum_{K^R_{so'}(t,T_P) <= K^R_{so}(t,T_P)} K^R_{so'}(t, T_P), \tag{10}$$

where $o, o' \in O$ We consider the future popularity of objects can be as a result of any of the three following scenarios:-

- It was popular in the past, its recent link gain score was high, and its total current rank score is also high, i.e. $(R^P_{so}(t, T_P)\&R^C_{so}(t)\&R^R_{so}(t, T_P))$.
- It was not popular in the past, but during the recent time it gained many links, and its current rank is high, i.e., $(\overline{R^P_{so}(t, T_P)}\&R^C_{so}(t)\&R^R_{so}(t, T_P))$.
- Its rank score was not high neither in the past nor in the current time, but its rank score is high during the recent time, i.e., $(\overline{R^P_{so}(t, T_P)}\&\overline{R^C_{so}(t)}\&R^R_{so}(t, T_P))$.

Since these three conditions are the most dominant and crucial for future popularity of an object, we combined all the three factors as follows:

$$\begin{aligned} P^F_s(o, T_P) = max\{&min\{R^P_{so}(t, T_P), R^C_{so}(t), R^R_{so}(t, T_P)\}, \\ &min\{\overline{R^P_{so}(t, T_P)}, R^C_{so}(t), R^R_{so}(t, T_P)\}, \\ &min\{\overline{R^P_{so}(t, T_P)}, R^C_{so}(t), R^R_{so}(t, T_P)\}\} \end{aligned} \tag{11}$$

TABLE 1. **Information about the processed data.**

Data set	Users	Objects	Links
Movielens 20M	10, 000	16, 433	1.2×10^6
Digg	10, 000	3, 553	2.1×10^5

where, $P^F_s(o, T_P)$ is the predicted score of object's popularity in a time after t. $\overline{R^P_{so}(t, T_P)} = 1 - R^P_{so}(t, T_P)$ and $\overline{R^C_{so}(t)} = 1 - R^C_{so}(t)$. We call this model "Non-Param" for ease of representation when plotting figures.

D. DATA

To test the accuracy of the predictor, we have used a movie rating and news item likes dataset.

- **Movielens 20M** contains about 20 million rating records of 27, 278 movies rated by 138493 users between 09 January,1995 and 31 March 2015 [58], [59]. Each user has rated a movie from 1 to 5 (1 is the worst and 5, the best). In this study, we considered only positive ratings (higher than 2). We randomly selected 10, 000 unique users and all the movies rated by them. Further, we have changed the time when the user rated the movie into days, the day we have counted from the time we have data (9 January, 1995). In consequence the event time in the system will be day 0 and next event time will be day 1 and so on.
- **Digg voting** contains 3, 553 news items where 139, 409 users have voted or digged (3, 018, 196 links) for the news stories from 31 January 2009 to 5 July 2009 [60], [61]. We have randomly sampled 10, 000 users and all their digged news. We converted the time into hours, since the start time.

The temporal distributions of the objects in the two data sets are shown in Fig. 2, depicting their final degree distribution in the bottom row. The information about the final processed data can be found in Table 1.

E. EVALUATION METRICS

We have adopted the following evaluation metrics to measure the accuracy of our proposed model: *Temporal Novelty* (TN_k), *Area Under receiving operating Characteristic* (AUC_k) also known as ROC [62], precision(P_k) and Kendal's rank correlation Tau(τ).

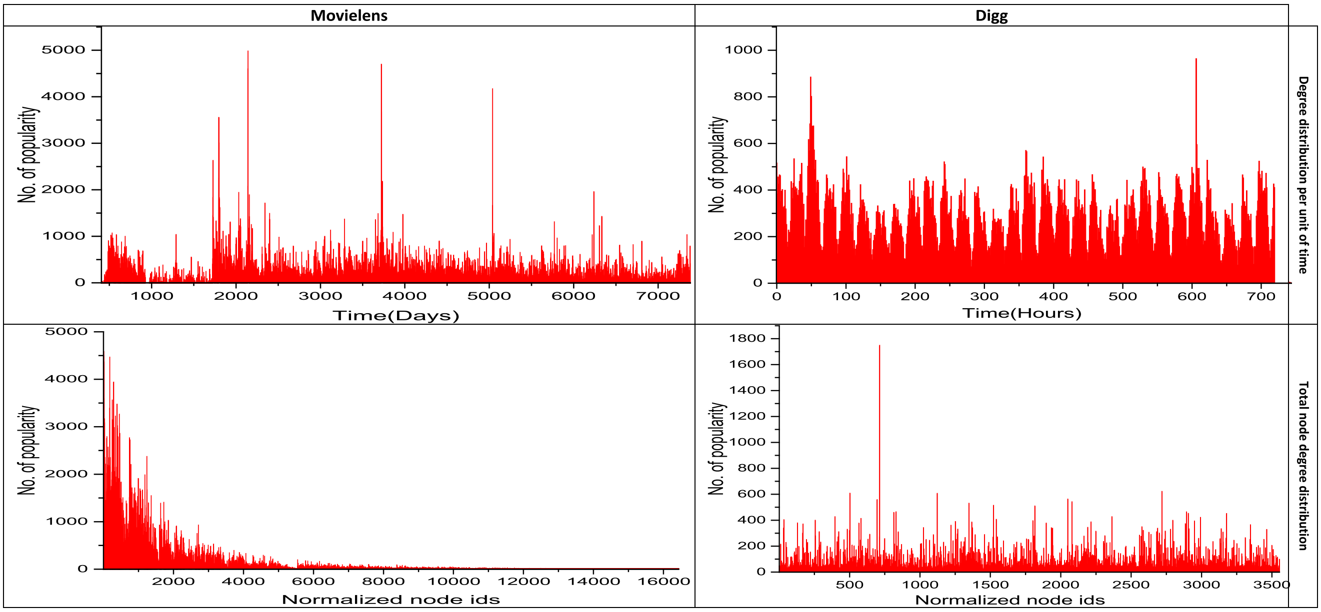

FIGURE 2. The nature of the data used in our experiment is as follows: First top row shows the temporal occurrence of link formation per unit of time(days for Movielens and hours for Digg). The second row shows the total degree distribution histogram.

- *Temporal Novelty(TN_k)* measures the ability of a predictor to rank 'new object' in top k list that was not popular during recent past time window but during future time window T_F they gained popularity. Let $R_k^{\Delta t}$ denote the number of new objects (that were not in top rank by popularity gain during recent time window T_P) in top k of the real list. And $E_k^{\Delta t}$ denotes the number of the new objects correctly predicted by our model in the top k ranking list. Then the temporal novelty (TN_k) score is given by-

$$TN_k = \frac{E_k^{\Delta t}}{R_k^{\Delta t}}, \tag{12}$$

- *AUC* measures the importance of the relative position of top k objects in the predicted and ranked lists. It selects the top k objects from the real list as a benchmark and compares their rank scores with the top k objects in the predicted list. Let $s_p \in L_p$ and $s_r \in L_r$ be the scores of an object in predicted lists respectively, then *AUC* can be calculated by:

$$AUC = \frac{\sum_{s_p \in L_p} \sum_{s_r \in L_r} I(s_p, s_r)}{|L_p|\,|L_r|} \; where, \tag{13}$$

$$I(s_p, s_r) = \begin{cases} 0, & \text{if } s_p > s_r, \\ 0.5, & \text{if } s_p = s_r, \\ 1, & \text{if } s_p < s_r. \end{cases} \tag{14}$$

- *Precision* is defined as the fraction of objects listed in the top k rankings of the predicted and real ranking lists [63] and is given by:

$$P_k = \frac{D_k}{k}, \tag{15}$$

where D_k is the number of common objects in the top k of both predicted and real ranking lists. $P_k \in [0, 1]$. The higher value of P_k, the better precision of prediction.

- *Novelty(Q_k)* measures the ability of a predictor to rank 'new objects' in the top k position which were not in the top k positions in the past. Let R_k denotes the number of 'new objects' in the top k position of the real list and E_K the number of 'new objects' correctly predicted by our model in the top k ranking list, then the novelty score is measured by:

$$Q_k = \frac{E_k}{R_k}, \tag{16}$$

- *Kendal's Tau(τ)* measures the correlation between predicted and real rankings. It varies between -1 and $+1$. $\tau = 1$ when predicted and real (actual) are identical, $\tau = 0$ when both ranking are independent and $\tau = -1$ shows they perfectly disagree. It can be given as-

$$\tau = \frac{C - D}{\sqrt{(C + D - N_{tp})}\sqrt{(C + D - N_{tr})}}, \tag{17}$$

where C is the number of concordant pairs and D is the number of discordant pairs. N_{tp} is the number of ties in predicted list and N_{tr} number of ties in real list.

IV. ANALYSIS AND RESULTS

This study aims to predict popularity of objects according to its link gain during a future time window, T_F (refer to Fig. 1). We converted this popularity prediction problem into a future link gain prediction problem in a bipartite network consisting two sets of nodes: users (U) and objects (O). A link from u to o ($u \rightarrow o$) will be formed, when user u uses object o. Likewise other prediction problems, we need to have a training set and a validation/test set. To evaluate the performance of our model (predictor), we have selected 10 random time t for training and testing purpose. Selection of random time t is considered in such a way that the predictor has enough historical information. Since predictors are based on objects' history, only objects with at least one link are selected. After picking random times we consider the future time window T_F. At time $t + T_F$ we calculate true top objects against which we measure accuracy of our proposed method. We evaluated the performance of our proposed model using a combination of three parameters: T_P (recent time window), T_F (future time window for prediction), and k (top-k ranking). The model was tested using various scenarios in which one or two of these parameters were varied while the others were held constant. Additionally, in the case of TBP, we performed a parameter search for γ by iterating over the range of [0, 1] with two decimal places precision, and selecting the value that yielded the highest precision.

A. VARYING LIST SIZE (k) AND FIXED PAST T_P AND FUTURE T_F TIME WINDOWS

In this section, we examine the sensitivity of our proposed non-parametric model to the variable k, based on the percentage of the selected list compared to the full list, while fixing the T_P and T_F. We selected the past and future time windows as 30 days for the Movielens dataset, and 10 hours for the Digg dataset. Fig.,3 compares the performance of our model using the evaluation metrics (i.e. AUC, Tau, and Temporal Novelty). The X-axis represents the size as a percentage of the whole list. As shown in Fig.,3, the AUC decreases as the list size increases for both datasets and temporal novelty is greatly affected by the list size. For the Digg dataset, it is zero, and for the Movielens dataset, it is very low. In terms of precision, after around 40%, it achieves 100% accuracy. In terms of novelty, it shows steady performance after $2 - 5\%$ for both datasets. We can conclude that our non-parametric model is efficient as it is computationally efficient to use a large portion of data, which produces maximum accuracy.

B. VARYING FUTURE TIME (T_F), FIXED PAST TIME (T_P) AND FIXED SIZE k

To set an appropriate fixed past time window (T_P), characteristics of datasets such as their evolution rate (the number of added nodes and links per unit of time), were considered. As the movie rating process is slower than the Digg voting process, we set a long period of 90 days as T_P for the Movielens, but 5 hours for Digg dataset. We evaluated the predictor for varying future time lengths from day 1 to 500 days in the case of Movielens and up to 50 hours for Digg dataset. The X-axis represents the time and Y-axis shows the accuracy results based on different evaluation metrics.

From Fig. 4 we have following results:

- In a **AUC** analysis for Movielens data, we found non-param performs same as TBP. In case of Digg, dataset non-param outperforms all others.
- In the analysis of **precision**(P_{100}), we found that TBP outperforms the Non-Param for the Movielens dataset, but not by a large margin. In the case of the Digg dataset, the performance of TBP is equal to that of Non-Param.
- In the analysis of **novelty** (Q_{100}), we found that TBP outperforms for the Movielens dataset, but in the case of the Digg dataset, the non-parametric method performs similarly to TBP. On the other hand, PA has no performance at all.
- In the analysis of **Temporal Novelty** (TN_{100}), we found that for the Movielens dataset, the non-parametric method outperforms others, but in the case of the Digg dataset, it has no performance at all.

C. VARYING FUTURE (T_F) AND PAST (T_P) TIME WINDOWS AND FIXED k SIZE

In Fig.,5, we have shown the results of varying past and future time windows (T_P and T_F) for all the evaluation metrics. The X-axis represents T_F and the Y-axis represents T_P, with the X-Y plane displaying every possible combination of past and future time windows within this range. We considered the top 100 objects ($k = 100$) for performance calculation, except for rank correlation tau(τ). The detailed analysis is as follows:

1) HEAT MAP ANALYSIS FOR MOVIELENS DATA SET

As shown in the first column of Fig.,5 for the Movielens data, we have found the following results:

- In the *AUC* analysis, there is a very small area in the (T_F,T_P) plane where AUC values are low. The blue dots are very few and the majority of the space shows high accuracy.
- In the *precision(P)* analysis, there is a very small area in the (T_F,T_P) plane where precision is low. There are many options for selecting a past time window that result in good precision.
- In the *novelty (Q)* analysis, as shown in Fig.,5, there are many (T_F,T_P) spaces for which accuracy is higher for all the evaluation metrics used: Precision, Novelty, AUC and tau(τ). Only for very short past time windows T_P, the performance is lower. Therefore, near the X and Y-axis, there is a density of blue dots which indicate lower accuracy. This means that on Movielens, novelty prediction is not very sensitive to the selection of past time window T_P.
- In the *Temporal Novelty (TN)* analysis, the region of lower accuracy is smaller but larger than in the case of novelty and precision. In most of the (T_F,T_P) space, the accuracy is similar. The region of high accuracy is also small. This suggests that it is sensitive to the selection of the past time window T_P.
- In the *rank correlation* (τ) analysis, we have found that it is similar to novelty, precision, and AUC, showing a higher density of red dots. This implies that there is a larger area of good accuracy in the (T_F,T_P) plane.

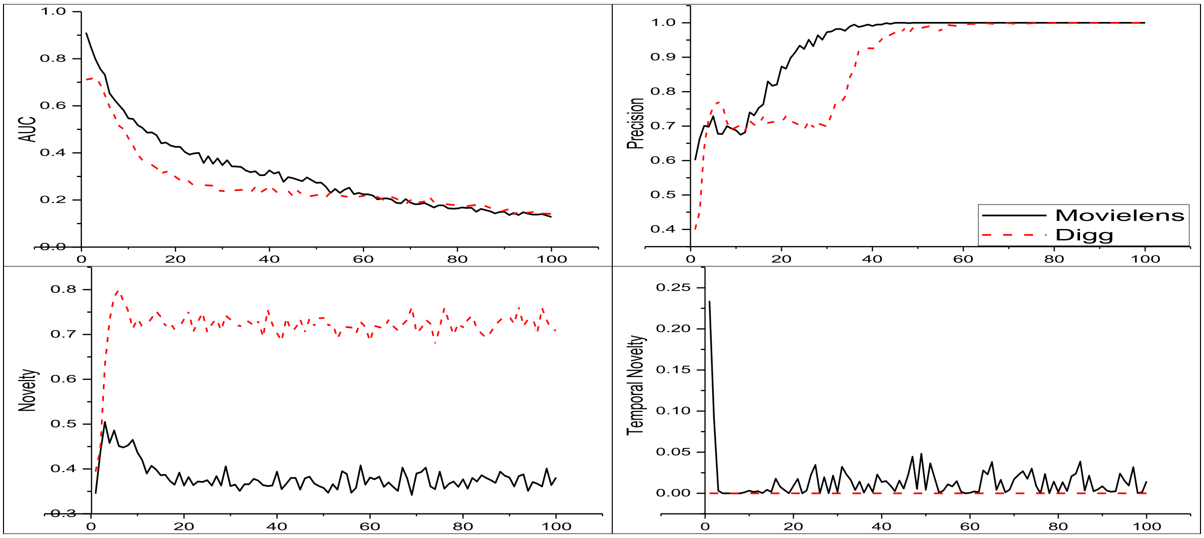

FIGURE 3. The sensitivity of the model was evaluated for varying top k values (number of selected objects for comparing predicted and real ranking lists) while keeping T_P and T_F fixed. In this scenario, the past and future time windows, T_P and T_F, were fixed at 30 days for the Movielens dataset and 10 hours for the Digg dataset. The X-axis represents the top k node list size as a percentage of the whole list.

2) HEAT MAP ANALYSIS FOR DIGG DATA SET

As shown in Fig.,5, in the second column from the left, for the Digg dataset, we have found the following results. To plot the heat map for the Digg dataset, we have considered the past and future time windows from the first hour to the 50th hour.

- In the *AUC* analysis, the area of higher accuracy is larger than in the cases of precision and novelty, as described above. However, it is concentrated around less than 50 hours T_F and less than 10 hours T_P.
- In the *precision(P)* analysis, lower past time windows help in predicting both long-term and short-term accuracy. However, after 30 hours (T_F) the prediction accuracy decreases. A similar effect is observed for past time window selection as well. In Fig.,5 on the Y-axis, past time windows (T_P) less than 10 hours give good results, while after 10 hours the performance decreases.
- In the *novelty (Q)* analysis, lower past time windows help in predicting both long-term and short-term trend prediction. However, after 30 hours, the prediction accuracy is not good. The heat map in Fig.,5 shows that as T_P increases the results become more blue, indicating lower accuracy. This suggests that our model is sensitive to past time window selection on this dataset.
- In the *Temporal Novelty* analysis, our model has no ability to make predictions for new items that did not gain popularity in the past time window.
- In the rank correlation *Tau(τ)* analysis, performance is not significantly affected by T_F, but it is affected by the past time window after 15 hours. As the past time window increases, the accuracy of the model decreases.

D. RANK CORRELATION WITH THE THREE COMPONENTS OF OUR MODEL

In Fig.6, we have plotted the rank correlation between the different components of our proposed model. Additionally, we have also plotted the correlation between the different components of our model and the actual future links. To do this, we used the Movielens dataset and calculated the rank correlation at every time step up to 500 days. At every time step, the past and future time window lengths are the same. We have found that $\overline{R^C_{so}(t, T_P)}\&R^R_s(t, T_P)$ and $R^C_{so}(t, T_P)\&\overline{R^R_s(t, T_P)}$ are negatively correlated. We also found that correlation between $R^P_{so}(t, T_P)\&\overline{R^R_s(t, T_P)}$ is the highest and it decreases as the time window increases. The reason is that the component $R^R_s(t, T_P)$ is completely dependent on time window size. The rest of the correlation shows a positive correlation with other components as shown in Fig. 6.

E. TIME COMPLEXITY ANALYSIS

In this section, we have evaluated the time complexity of our model and compared it to the benchmark models. Suppose there are T snapshots of the graph, and each snapshot has n nodes and m edges. As we are calculating the Big-O notation, we will consider the worst-case scenario.

- The **Barabasi-Albert (BA)** model is very simple, and it's calculated by considering only the current snapshot, so its time complexity is $O(mn)$.
- The **TBP** model is complex and requires the learning of the parameter γ. Thus, the total time complexity of the model, $s_o(t) = \sum_u A_{uo}(t)\exp(\gamma(T_{uo} - t))$, including the learning of γ using stochastic gradient descent (SGD) in a dynamic setting with T snapshots of the graph dynamic $G(t)$, is $O(Tn^2 + Tmn)$, where m is the number of edges in the graph, and n is the number of nodes. The time complexity of calculating the score of the model, $s_o(t)$, for each snapshot is $O(n^2)$. This is because the model involves a summation over all nodes u, and for each node, it involves the computation of the matrix

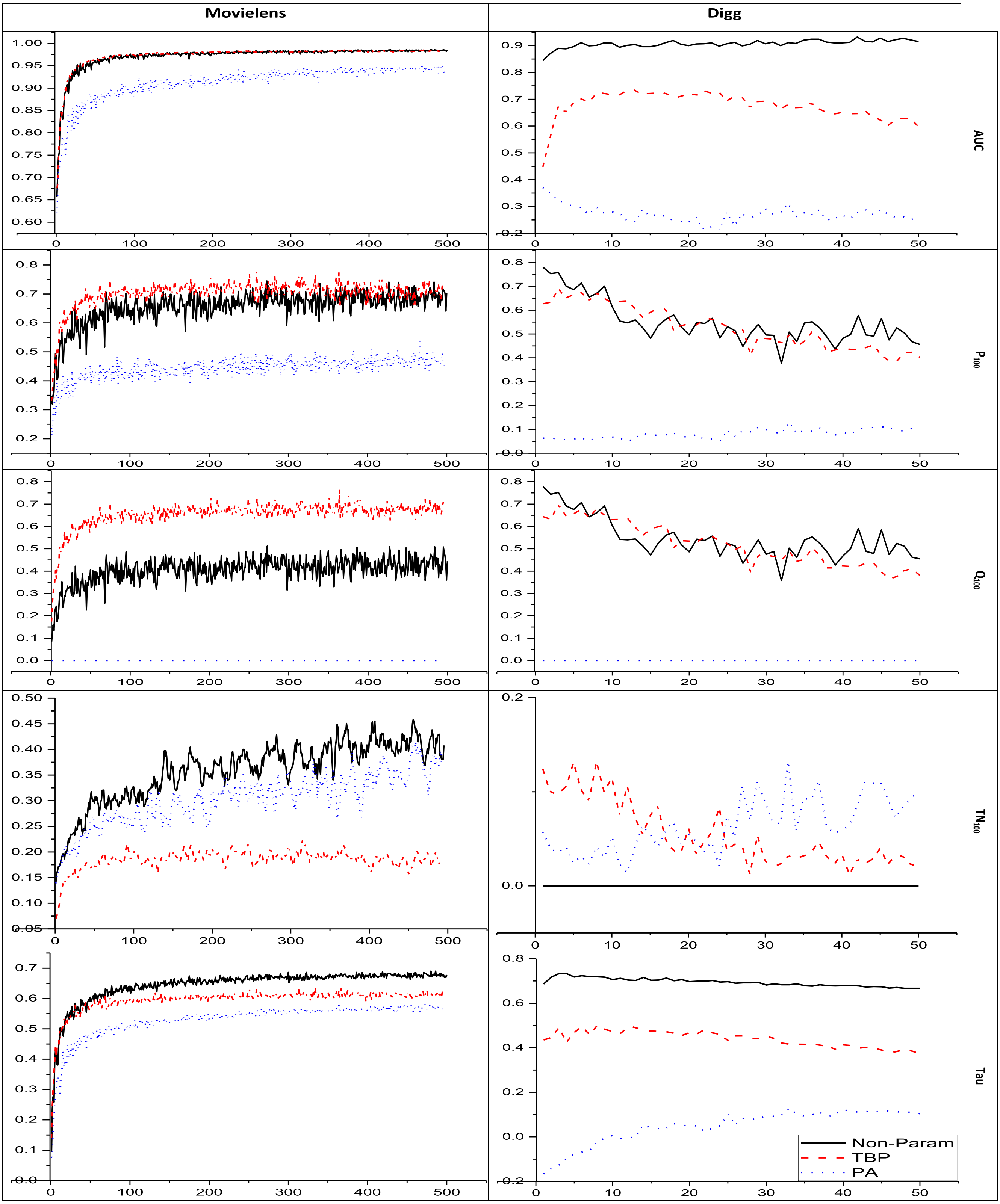

FIGURE 4. The sensitivity of the model's performance for fixed past and varying future time windows is shown on the X-axis. In this scenario, the past time window is fixed at 90 days for the Movielens dataset and 5 hours for the Digg dataset. The X-axis represents the time and the Y-axis represents the accuracy score, with a higher score being better. All scores lie between 0 and 1, except for the rank correlation tau (τ) which lies between −1 and 1. The time period is up to 500 days for the Movielens dataset and in hours for the Digg dataset, up to 50 hours. The red line with the dash (—) represents *TBP*, the dotted (...) blue line represents *PA*, and the solid black line represents the proposed model *Non − Param*.

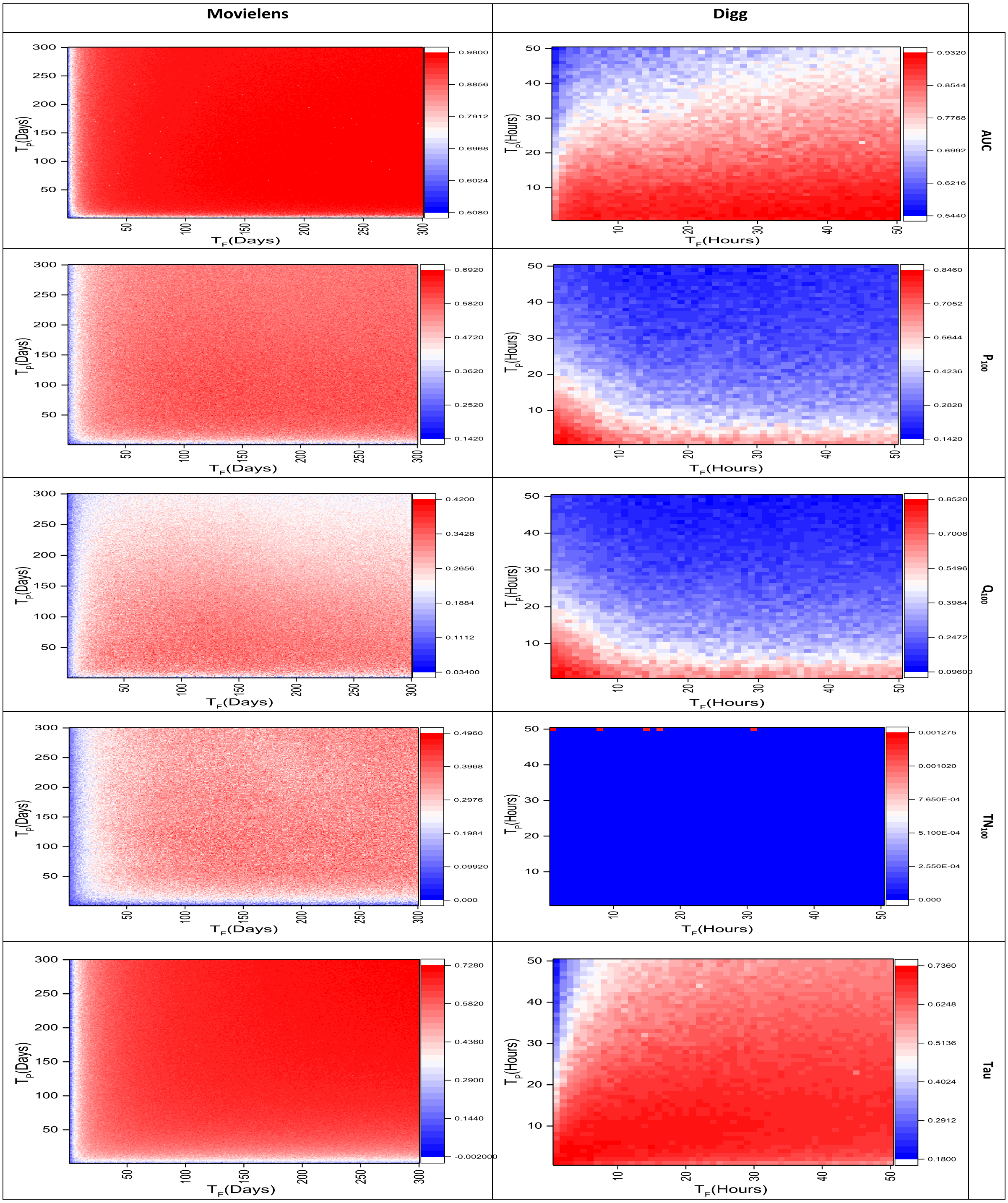

FIGURE 5. The sensitivity of the proposed model is depicted as a heat map, which plots the variation in past and future time windows, creating a (T_F, T_P) plane. For the Movielens dataset, the time is in days up to 300 days and for the Digg dataset, the time is in hours up to 50 hours.

product $A_{uo}(t)$ and the exponential function $\exp(\gamma(T_{uo} - t))$. Both of these operations have a time complexity of O(n), and since they are performed for each node, the overall time complexity is O(n^2). This needs to be done for T snapshots, so the time complexity for this would be O(Tn^2). For learning the parameter γ using SGD, the time complexity for each iteration is O(mn), as we are iterating over the edges in the graph. The number of iterations required to converge to a solution depends on the specific optimization algorithm and the

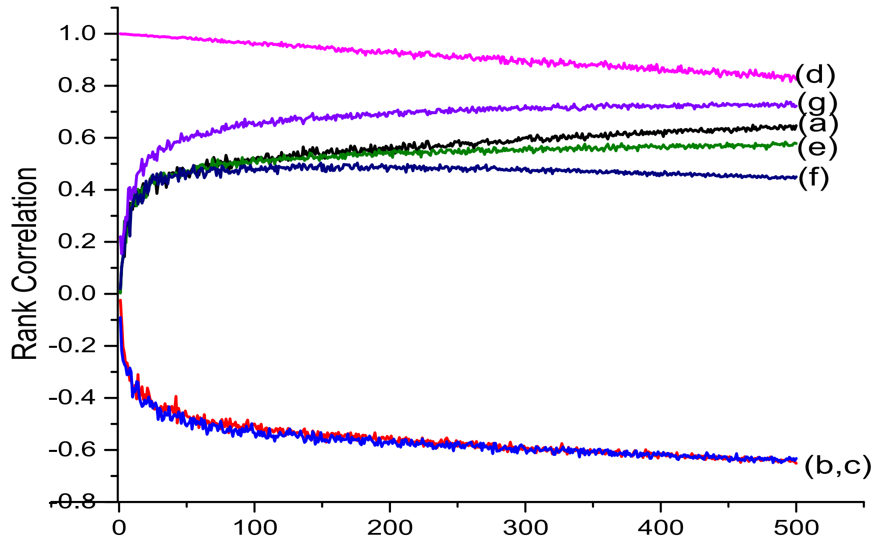

FIGURE 6. The rank correlation among the components used in the proposed model is analyzed by considering varying past and future time windows up to 500 days. At every time step, the past and future time window sizes are kept equal. The label (a) is for correlation between $R^{C}_{SO}(t, T_P)\&R^{R}_{S}(t, T_P)$; (b) is for correlation between $R^{C}_{SO}(t, T_P)\&R^{R}_{S}(t, T_P)$; (c) is for correlation between $R^{C}_{SO}(t, T_P)\&R^{R}_{S}(t, T_P)$; (d) is for correlation between $R^{P}_{SO}(t, T_P)\&R^{R}_{S}(t, T_P)$; (e) is for correlation between $R^{C}_{SO}(t, T_P)\&Actual$; (f) is for correlation between $R^{P}_{SO}(t, T_P)\&Actual$ and (g) is for correlation between $R^{R}_{S}(t, T_P)\&Actual$.

TABLE 2. Time complexity analysis with benchmark models.

Model	Time complexity	(m>>n)	(m<<n)
BA	O(mn)	O(mn)	O(mn)
TBP	Tn^2 + Tmn	O(Tmn)	O(Tn^2)
Non-param	mn+n+n^2	O(mn)	O(n^2)

characteristics of the dataset. Thus, the time complexity for parameter learning is O(Tmn). Hence, the total time complexity of the model is O($Tn^2 + Tmn$).

- The proposed **non-parametric** model has four simple steps. The first step, similar to the BA model, involves calculating three vectors and has a time complexity of O(mn), where m is the number of edges in the graph and n is the number of nodes. The second step involves normalizing the three vectors, which has a time complexity of O(n). The third step involves obtaining the corresponding cumulative distribution function (CDF), which has a time complexity of O(n^2). The fourth step involves assigning a rank score using Eq.(11), which has a time complexity of O(n). Therefore, the overall time complexity of the proposed non-parametric model is O($mn + n + n^2$).

From Table 2 analysis we can say that considering performance for dynamic graphs our proposed model has best time complexity. BA model has the best in all the cases but it doesn't have better performance.

V. CONCLUSION

Our proposed model is based on this uncertainty and it considers both past and recent past popularity of items. We have tested our model on two real-world datasets: Movielens and Digg. We have adopted various evaluation metrics such as AUC, Temporal Novelty,Precision, and Kendal's rank correlation Tau to measure the accuracy of our proposed model. We have also analyzed the sensitivity of our model to different parameters such as past and future time window, and the number of items on the top of the ranking list. Our results show that our model outperforms the other existing models in most of the cases, and it is computationally efficient for large datasets. Additionally, we have also found that the proposed model is sensitive to the past time window selection and it is better in predicting the long-term popularity of items.

We solved the problem of item popularity prediction using a rank-based approach, as it is widely applied to solve the information overload problem. We tested our proposed model on two real-world datasets and found that our model has a competent performance on five indices (i.e. AUC, precision, novelty, temporal novelty and Kendall's rank correlation (τ)). Our model only considers a node's collective link gain in the recent past as well as the total link gain. Our results show that the model has a competent accuracy compared to benchmark models. Additionally, our model is simple and does not require the learning of any hyperparameters, making it a suitable choice for real-time applications, particularly in scenarios where data size is very large.

Our results also indicate that the future popularity gain depends on three attributes: 1) recent popularity gain, i.e rating or link-gain in recent past time window; 2) link-gain during past time, i.e before recent time-window and 3) up to current time. The reason for considering past time-window popularity is that some content may gain popularity multiple times. Therefore, according to our intuitive model, even if a node/item has not received links during the recent time, if it was popular in the past, it can gain popularity in the future. We found that considering these three time-windows improves the accuracy of predicting future popularity gain.

As a next step for the proposed heuristics-based node ranking model, it would be beneficial to consider incorporating various features into the model. This includes incorporating various types of node-level features such as various centrality measures or community membership, various types of edge-level features like weight or direction, and different edge types such as directed or undirected edges. These features can potentially improve the predictive performance of the model. Additionally, future research can include experimenting with different heuristics and ensemble techniques to find the optimal combination of heuristics for different types of networks. Furthermore, evaluating the proposed model on a diverse set of real-world dynamic networks and comparing its performance with other state-of-the-art models can also be considered as a future work.

REFERENCES

[1] A. Sinha, Z. Shen, Y. Song, H. Ma, D. Eide, B.-J. Hsu, and K. Wang, "An overview of Microsoft academic service (MAS) and applications," in *Proc. 24th Int. Conf. World Wide Web*, May 2015, pp. 243–246.

[2] A. Khan, J. P. Li, N. Ahmad, S. Sethi, A. U. Haq, S. H. Patel, and S. Rahim, "Predicting emerging trends on social media by modeling it as temporal bipartite networks," *IEEE Access*, vol. 8, pp. 39635–39646, 2020.

[3] A. D. Wade, K. Wang, Y. Sun, and A. Gulli, "WSDM cup 2016: Entity ranking challenge," in *Proc. 9th ACM Int. Conf. Web Search Data Mining*, Feb. 2016, pp. 593–594.

[4] H. Liao, M. S. Mariani, M. Medo, Y.-C. Zhang, and M.-Y. Zhou, "Ranking in evolving complex networks," *Phys. Rep.*, vol. 689, pp. 1–54, May 2017.

[5] A. Zeng, S. Gualdi, M. Medo, and Y.-C. Zhang, "Trend prediction in temporal bipartite networks: The case of Movielens, Netflix, and Digg," *Adv. Complex Syst.*, vol. 16, nos. 4–5, Aug. 2013, Art. no. 1350024.

[6] M. Medo, "Network-based information filtering algorithms: Ranking and recommendation," in *Dynamics on and of Complex Networks*, vol. 2. Cham, Switzerland: Springer, 2013, pp. 315–334.

[7] L. Lü, M. Medo, C. H. Yeung, Y.-C. Zhang, Z.-K. Zhang, and T. Zhou, "Recommender systems," *Phys. Rep.*, vol. 519, no. 1, pp. 1–49, 2012.

[8] X. He, M. Gao, M.-Y. Kan, and D. Wang, "Birank: Towards ranking on bipartite graphs," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 1, pp. 57–71, Jan. 2017.

[9] K. Abbas, M. Shang, X. Luo, and A. Abbasi, "Emerging trends in evolving networks: Recent behaviour dominant and non-dominant model," *Phys. A, Stat. Mech. Appl.*, vol. 484, pp. 506–515, Oct. 2017.

[10] K. Abbas, M. Shang, A. Abbasi, X. Luo, J. J. Xu, and Y.-X. Zhang, "Popularity and novelty dynamics in evolving networks," *Sci. Rep.*, vol. 8, no. 1, pp. 1–10, Apr. 2018.

[11] H. C. X. Li, "Recommendation as link prediction in bipartite graphs: A graph kernel-based machine learning approach," *Decis. Support Syst.*, vol. 54, pp. 880–890, Jan. 2013.

[12] A. Ozcan and S. G. Oguducu, "Link prediction in evolving heterogeneous networks using the NARX neural networks," *Knowl. Inf. Syst.*, vol. 55, no. 2, pp. 333–360, 2018.

[13] A. Ozcan and S. G. Oguducu, "Multivariate time series link prediction for evolving heterogeneous network," *Int. J. Inf. Technol. Decis. Making*, vol. 18, no. 1, pp. 241–286, 2019.

[14] M. Gayberi and S. G. Oguducu, "Popularity prediction of posts in social networks based on user, post and image features," in *Proc. 11th Int. Conf. Manage. Digit. EcoSyst.*, Nov. 2019, pp. 9–15.

[15] T. Preis, H. S. Moat, and H. E. Stanley, "Quantifying trading behavior in financial markets using Google Trends," *Sci. Rep.*, vol. 3, no. 1, pp. 1–6, Apr. 2013.

[16] C. H. Hommes, "Modeling the stylized facts in finance through simple nonlinear adaptive systems," *Proc. Nat. Acad. Sci. USA*, vol. 99, no. 3, pp. 7221–7228, May 2002.

[17] J.-P. Bouchaud, A. Matacz, and M. Potters, "Leverage effect in financial markets: The retarded volatility model," *Phys. Rev. Lett.*, vol. 87, no. 22, Nov. 2001.

[18] R. L. Axtell, "Zipf distribution of U.S. firm sizes," *Science*, vol. 293, no. 5536, pp. 1818–1820, Sep. 2001.

[19] G. King, "Ensuring the data-rich future of the social sciences," *Science*, vol. 331, no. 6018, pp. 719–721, 2011.

[20] M. Perc, "Evolution of the most common english words and phrases over the centuries," *J. Roy. Soc. Interface*, vol. 9, no. 77, pp. 3323–3328, Dec. 2012.

[21] T. Iwata, A. Shah, and Z. Ghahramani, "Discovering latent influence in online social activities via shared cascade Poisson processes," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2013, pp. 266–274.

[22] H. Zhu, X. Wang, and J.-Y. Zhu, "Effect of aging on network structure," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 68, no. 5, Nov. 2003, Art. no. 056121.

[23] G. O. Mohler, M. B. Short, P. J. Brantingham, F. P. Schoenberg, and G. E. Tita, "Self-exciting point process modeling of crime," *J. Amer. Statist. Assoc.*, vol. 106, no. 493, pp. 100–108, 2011.

[24] Q. Zhao, M. A. Erdogdu, H. Y. He, A. Rajaraman, and J. Leskovec, "SEISMIC: A self-exciting point process model for predicting tweet popularity," in *Proc. 21st ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2015, pp. 1513–1522, doi: 10.1145/2783258.2783401.

[25] S. Gao, J. Ma, and Z. Chen, "Modeling and predicting retweeting dynamics on microblogging platforms," in *Proc. 8th ACM Int. Conf. Web Search Data Mining*, Feb. 2015, pp. 107–116.

[26] P. D. B. Parolo, R. K. Pan, R. Ghosh, B. A. Huberman, K. Kaski, and S. Fortunato, "Attention decay in science," *J. Informetrics*, vol. 9, no. 4, pp. 734–745, Oct. 2015.

[27] D. Wang, C. Song, and A.-L. Barabási, "Quantifying long-term scientific impact," *Science*, vol. 342, no. 6154, pp. 127–132, 2013.

[28] H. Shen, D. Wang, C. Song, and A.-L. Barabási, "Modeling and predicting popularity dynamics via reinforced Poisson processes," in *Proc. 38th AAAI Conf. Artif. Intell.*, vol. 14, Jul. 2014, pp. 291–297.

[29] E. Bakshy, J. M. Hofman, W. A. Mason, and D. J. Watts, "Everyone's an influencer: Quantifying influence on Twitter," in *Proc. 4th ACM Int. Conf. Web Search Data Mining*, Feb. 2011, pp. 65–74, doi: 10.1145/1935826.1935845.

[30] T. Martin, J. M. Hofman, A. Sharma, A. Anderson, and D. J. Watts, "Exploring limits to prediction in complex social systems," in *Proc. 25th Int. Conf. World Wide Web*, Apr. 2016, pp. 683–694.

[31] G. Szabo and B. A. Huberman, "Predicting the popularity of online content," *Commun. ACM*, vol. 53, no. 8, pp. 80–88, 2010.

[32] J. Cheng, L. Adamic, P. A. Dow, J. M. Kleinberg, and J. Leskovec, "Can cascades be predicted?" in *Proc. 23rd Int. Conf. World Wide Web*, Apr. 2014, pp. 925–936.

[33] L. Weng, F. Menczer, and Y. Y. Ahn, "Virality prediction and community structure in social networks," *Sci. Rep.*, vol. 3, no. 1, p. 2522, 2013.

[34] R. A. Bentley, M. W. Hahn, and S. J. Shennan, "Random drift and culture change," *Proc. Roy. Soc. London. Ser. B, Biol. Sci.*, vol. 271, no. 1547, pp. 1443–1450, Jul. 2004.

[35] R. A. Bentley, M. Earls, and M. J. O'Brien, *I'll Have What She's Having: Mapping Social Behavior*. Cambridge, MA, USA: MIT Press, 2011.

[36] C. Candia, C. Jara-Figueroa, C. Rodriguez-Sickert, A.-L. Barabási, and C. A. Hidalgo, "The universal decay of collective memory and attention," *Nature Hum. Behav.*, vol. 3, no. 1, pp. 82–91, Dec. 2018.

[37] J. Cheng, L. A. Adamic, J. M. Kleinberg, and J. Leskovec, "Do cascades recur?" in *Proc. 25th Int. Conf. World Wide Web*, Apr. 2016, pp. 671–681.

[38] F. Wu and B. A. Huberman, "Novelty and collective attention," *Proc. Nat. Acad. Sci. USA*, vol. 104, no. 45, pp. 17599–17601, 2007.

[39] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," *Science*, vol. 359, pp. 1146–1151, May 2018.
[40] Y. Zhou, A. Zeng, and W.-H. Wang, "Temporal effects in trend prediction: Identifying the most popular nodes in the future," *PLoS ONE*, vol. 10, no. 3, Mar. 2015, Art. no. e0120735.
[41] J. Liebig and A. Rao, "Predicting item popularity: Analysing local clustering behaviour of users," *Phys. A, Statist. Mech. Appl.*, vol. 442, pp. 523–531, Jan. 2016.
[42] S. Lin, X. Kong, and P. S. Yu, "Predicting trends in social networks via dynamic activeness model," in *Proc. 22nd ACM Int. Conf. Inf. Knowl. Manag.*, 2013, pp. 1661–1666.
[43] S. Gao, J. Ma, and Z. Chen, "Effective and effortless features for popularity prediction in microblogging network," in *Proc. 23rd Int. Conf. World Wide Web*, Apr. 2014, pp. 269–270.
[44] S. Mishra, "Bridging models for popularity prediction on social media," in *Proc. 12th ACM Int. Conf. Web Search Data Mining*, Jan. 2019, pp. 810–811.
[45] F. Nsour and H. Sayama, "Hot-get-richer network growth model," in *Proc. Int. Conf. Complex Netw. Their Appl.* in Studies in Computational Intelligence, vol. 944, R. M. Benito, C. Cherifi, H. Cherifi, E. Moro, L. M. Rocha, and M. Sales-Pardo, Eds. Madrid, Spain: Springer, 2020, pp. 532–543.
[46] S. Pongnumkul and K. Motohashi, "A bipartite fitness model for online music streaming services," *Phys. A, Stat. Mech. Appl.*, vol. 490, pp. 1125–1137, Jan. 2018.
[47] A. Chandra, H. Garg, and A. Maiti, "A general growth model for online emerging user–object bipartite networks," *Phys. A, Stat. Mech. Appl.*, vol. 517, pp. 370–384, Mar. 2019.
[48] Y. Zhou, Q. Li, X. Yang, and H. Cheng, "Predicting the popularity of scientific publications by an age-based diffusion model," *J. Informetrics*, vol. 15, no. 4, Nov. 2021, Art. no. 101177.
[49] M. Kitsak, "Identification of influential spreaders in complex networks," *Nature Phys.*, vol. 6, pp. 888–893, Aug. 2010.
[50] L.-L. Ma, C. Ma, H.-F. Zhang, and B.-H. Wang, "Identifying influential spreaders in complex networks based on gravity formula," *Phys. A, Stat. Mech. Appl.*, vol. 451, pp. 205–212, Jul. 2016.
[51] F. Yang, R. Zhang, Z. Yang, R. Hu, M. Li, Y. Yuan, and K. Li, "Identifying the most influential spreaders in complex networks by an extended local K-shell sum," *Int. J. Modern Phys. C*, vol. 28, no. 1, Jan. 2017, Art. no. 1750014.
[52] L. Qiu, J. Zhang, and X. Tian, "Ranking influential nodes in complex networks based on local and global structures," *Int. J. Speech Technol.*, vol. 51, no. 7, pp. 4394–4407, Jul. 2021.
[53] D. Chen, L. Lü, M.-S. Shang, Y.-C. Zhang, and T. Zhou, "Identifying influential nodes in complex networks," *Phys. A, Statist. Mech. Appl.*, vol. 391, pp. 1777–1787, Feb. 2012.
[54] F. Radicchi, S. Fortunato, and C. Castellano, "Universality of citation distributions: Toward an objective measure of scientific impact," *Proc. Nat. Acad. Sci. USA*, vol. 105, no. 45, pp. 17268–17272, Nov. 2008.
[55] F. Radicchi and C. Castellano, "Rescaling citations of publications in physics," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 83, no. 4, Apr. 2011, Art. no. 046116.
[56] M. S. Mariani, M. Medo, and Y.-C. Zhang, "Ranking nodes in growing networks: When PageRank fails," *Sci. Rep.*, vol. 5, no. 1, Nov. 2015, Art. no. 16181.
[57] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, Oct. 1999.
[58] F. M. Harper and J. A. Konstan, "The movielens datasets: History and context," *ACM Trans. Interact. Intell. Syst.*, vol. 5, pp. 19:1–19:19, Dec. 2015.
[59] (2017). *Movielens 20 Million Ratings Dataset*. Accessed: Mar. 1, 2017. [Online]. Available: https://grouplens.org/datasets/movielens/20m/
[60] (2017). *Digg Votes Network Dataset KONECT*. Accessed: Mar. 1, 2017. [Online]. Available: http://konect.uni-koblenz.de/networks/digg-votes
[61] T. Hogg and K. Lerman, "Social dynamics of Digg," *EPJ Data Sci.*, vol. 1, no. 1, p. 5, 2012.
[62] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.
[63] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, "Evaluating collaborative filtering recommender systems," *ACM Trans. Inf. Syst.*, vol. 22, no. 1, pp. 5–53, 2004.

KHUSHNOOD ABBAS received the bachelor's and graduate degrees from Aligrah Muslim University, India, and the Ph.D. degree from the School of Computer Science and Technology, University of Electronic Science and Technology, in 2018. He is currently an Assistant Professor with Zhoukou Normal University, Henan, China. His current research interest includes data mining, specifically within the context of social and information networks.

MOHAMMAD KAMRUL HASAN (Senior Member, IEEE) received the Doctor of Philosophy (Ph.D.) degree in electrical and communication engineering from the Faculty of Engineering, International Islamic University, Malaysia, in 2016. He is currently a Senior Lecturer with the Faculty of Information Science and Technology, Center for Cyber Security, Universiti Kebangsaan Malaysia (UKM). He is specialized in elements pertaining to cutting-edge information-centric networks, such as computer networks, data communication and security, mobile network and privacy protection, cyber-physical systems, the Industrial IoT, transparent AI, and electric vehicle networks. He has published more than 150 indexed papers in ranked journals and conference proceedings. He is a member of the Institution of Engineering and Technology and the Internet Society. He is a certified Professional Technologist (P.Tech /Ts.) in Malaysia. He also served as the Chair for the IEEE Student Branch, from 2014 to 2016. He has actively participated in many events/workshops/training for the IEEE and IEEE humanity programs in Malaysia. He is an Editorial Member in many prestigious high-impact journals, such as IEEE, IET, Elsevier, Frontier, and MDPI, and the general chair, the co-chair, and a speaker of conferences and workshops for the shake of society and academy knowledge building, sharing, and learning. He has been contributing and working as a volunteer for under privileged people for the welfare of society.

ALIREZA ABBASI is currently the Director of Postgraduate Studies with the School of Engineering and IT, UNSW Canberra. His research interests include enhancing better functioning and performance of complex socio-technical systems, which primarily lies in the emerging field of social computing, and leveraging the capacity to collect and analyze data to reveal patterns of individual and group behavior.

SHI DONG received the M.S. degree in computer science from the University of Electronic and Technology of China, in 2009, and the Ph.D. degree in computer science from Southeast University. He was a Postdoctoral Researcher with the Huazhong University of Science and Technology and a Visiting Scholar with Washington University, St. Louis. He is currently a Distinguished Professor with Zhoukou Normal University. His current research interests include network management and network security.

TAHER M. GHAZAL received the B.Sc. degree in software engineering from Al Ain University, in 2011, the M.Sc. degree in information technology management from The British University in Dubai, which is associated with The University of Manchester and The University of Edinburgh, in 2013, the Ph.D. degree in IT/software engineering from Damascus University, in 2019, and the Ph.D. degree in information science and technology from Universiti Kebangsaan Malaysia, in 2023. With over a decade of extensive and diverse experience, he has served as an instructor, a tutor, a researcher, a teacher, an IT support/specialist engineer, and a business/systems analyst. He has worked in various departments, including the Department of Engineering, Department of Computer Science, Department of ICT, has worked as the Head of STEM and Innovation, and has also been involved in quality assurance, accreditation, and data analysis, in several governmental and private educational institutions under KHDA, Ministry of Education, and Ministry of Higher Education and Scientific Research, United Arab Emirates. His research interests include the IoT, IT, artificial intelligence, information systems, software engineering, web development, building information modeling, quality of education, management, big data, the quality of software, and project management. He is actively involved in community services in the projects and research field.

SITI NORUL HUDA SHEIKH ABDULLAH received the degree in computing from the University of Manchester Institute of Science and Technology, U.K., the master's degree in artificial intelligence from Universiti Kebangsaan Malaysia, and the Ph.D. degree in computer vision from the Faculty of Electrical Engineering, Universiti Teknologi Malaysia. While starting her career, she was involved in conducting national and international activities, such as Cyber Security Academia Malaysia (CSAM), FIRA, Asian Foundation, IMWU, DFRS, MIAMI, MACE, and IDB Alumni. After working as the Chairperson of the Center for Cyber Security, she is currently the Deputy Dean of Research and Innovation with the Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, and the Chairperson of CSAM. She has published four books titled *Pattern Recognition*, *Computational Intelligence in Data Science Application*, *Smart Prediction of Suspect's Serial Crime Location*, and *Optical Script Reader: Texture Binary Innovation*, and more than 50 journals and 100 conference manuscripts correspondingly. Her research interests include digital forensics, pattern recognition, and computer vision surveillance systems.

ASIF KHAN (Member, IEEE) received the B.Sc. (Hons.) and Master of Computer Science and Application (M.C.A.) degrees from Aligarh Muslim University, India, and the Ph.D. degree (Hons.) in computer science and technology from the University of Electronic Science and Technology of China (UESTC), China, in 2016. He was an Adjunct Faculty with the University of Bridgeport, USA, for the China Program, in Summer 2016. Previously, he was a Visiting Scholar of big data mining and application with the Chongqing Institute of Green and Intelligent Technology (CIGIT), Chinese Academy of Sciences, Chongqing, China. He was a Postdoctoral Scientific Research Fellow with UESTC. He is currently an Assistant Professor with Integral University, India. He is a contributor to many international journals with robotics and vision analyses about the contemporary world in his articles. His research interests include machine learning, robotics vision, and new ideas regarding vision-based information critical theoretical research. He was a recipient of the UESTC Academic Achievement Award and the Excellent Performance Award, in 2015 and 2016, respectively.

DABIAH ALBOANEEN received the M.Sc. degree (Hons.) in advanced computer networking and the Ph.D. degree in cloud computing and artificial intelligence from Glasgow Caledonian University, in 2013 and 2019, respectively. She is currently an Assistant Professor with the College of Science and Humanities, Imam Abdulrahman Bin Faisal University, Jubail. She was awarded a lot of prizes, such as the Best Master Project in Networking and Wireless Communication Prize by the Glasgow Caledonian University, in 2013, and 11 distinction awards from Saudi Cultural Bureau in London, since 2013.

FATIMA RAYAN AWAD AHMED received the B.S. degree in computer science from the Sudan University of Science and Technology, in 2004, and the M.S. and Ph.D. degrees in computer science from Al Neelain University, Sudan, in 2007 and 2012, respectively. In 2004, she joined the IT Department, Sudanese Company for Telecommunications, Sudan, as a Computer Programmer, where she analyzed, designed, and programmed a set of systems. She joined Prince Sattam Bin Abdulaziz University, Saudi Arabia, in 2013, as an Assistant Professor with the Information Systems Department, from 2013 to 2016, and has been with the Computer Science Department, since 2017. Her research interests include artificial intelligence, systems and algorithms analysis and design, web applications, and e-learning.

THOWIBA E. AHMED received the B.S. degree in computer science and statistics from Al Neelain University, Sudan, in 2004, and the M.S. and Ph.D. degrees in information technology from Al Neelain University, Sudan, in 2007 and 2012, respectively. She has been with the College of Science and Humanities, Imam Abdulrahman Bin Faisal University, Saudi Arabia, since 2013. She was an Assistant Professor with the Computer Science Department and the Head of the Computer Science Department, from 2016 to 2019. In 2004, she joined the National Highway Authority, Sudan, as a Programmer, and after that as the Information Network Supervisor. She also joined Al Neelain University as a collaborative Teaching Assistant, in 2004, and as a collaborative Assistant Professor, in 2012. In 2012, she joined the Emirates College of Science and Technology, Sudan, as an Assistant Professor and the Head of the Information Technology Department. Her research interests include data science, HCI, systems and algorithms analysis and design, web applications, and e-learning.

SHAYLA ISLAM (Member, IEEE) received the B.Sc. degree in computer science and engineering from International Islamic University Chittagong, Bangladesh, the M.Sc. degree from the Department of Electrical and Computer Engineering, International Islamic University Malaysia (IIUM), in 2012, and the Ph.D. degree in engineering from the Electrical and Computer Engineering (ECE) Department, IIUM, in 2016, under Malaysian International Scholarship (MIS). She is currently an Assistant Professor with UCSI University, Malaysia. She was awarded a Silver Medal for her research work with International Islamic University Malaysia. In consequence, she was also awarded a Young Scientist Award for the contribution of a research paper at the second International Conference on Green Computing and Engineering Technologies, 2016 (ICGCET'16), organized by the Department of Energy Technology, Aalborg University, Esbjerg, Denmark.