This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TNNLS.2023.3330100, IEEE Transactions on Neural Networks and Learning Systems

Multi-Agent-System-Based Attention Mechanism for Predicting Product Popularity: Handling Positive-Negative Diffusion Over Social Networks

Mincan Li, Zidong Wang, *Fellow, IEEE*, Kenli Li, and Xiangke Liao

***Abstract*—This paper is concerned with the prediction problem of product popularity under a social network (SN) with positive-negative diffusion (PND). Firstly, a PND model is proposed to enable the simulation of product diffusion, and three user states are defined. Secondly, an optimal and precise feature vector of every user is extracted through a multi-agent-system-based attention mechanism (MASAM) that is devised. The weight matrix shared in the mechanism of all agents is learned using a distributed learning algorithm provided in MASAM. Thirdly, an MAS model for product diffusion on SN is established based on the feature representations from MASAM. Rules for agent interaction during PND diffusion are suggested, which accelerate the simulation of information spread in SN. Finally, comprehensive experiments are conducted to verify the effectiveness and efficiency of the proposed models and algorithms in prediction and to compare their performance with baseline methods. Furthermore, a case study is provided to illustrate the applicability and extendibility of the developed algorithm.**

***Index Terms*—Information diffusion, graph attention network, multi-agent systems, social networks, popularity prediction.**

I. Introduction

For several decades, social networks (SNs) have been considered to be effective platforms for product promotion, advertisement marketing, and information diffusion. Various social applications have been used to promote online SNs, and a large proportion of research has been focused on SN analysis. However, it should not be ignored that human activities follow and influence online SNs all the time and everywhere. Technologies such as data mining, relationship analysis, community detection, link prediction, and node classification have been developed by studying the information and structure of SNs [32].

Numerous applications and analysis models have been proposed on SNs to promote economic development [5], [6], [20], [22]. For instance, a vehicle tracking system can be established by integrating SN services. This system can be used for traffic management, vehicle tracking, anti-theft measures, and traffic routing with navigation. Real-data mining on SNs has led to the creation of a framework for predicting and warning people about disasters, which has the potential to be developed into a fully-fledged application for family systems. Another proposed system is an analysis tool for procurement and project management in construction projects that relies on the analysis of dependency information and partnering relationships in SNs. Also, from the modeling perspective, SNs are a special class of multi-agent systems (MASs) [22], which can be identified via data-driven approaches with applications to manufacturing systems [46] and cyber physical systems [47].

The prediction problem, which includes link prediction and popularity prediction, is one of the most popular topics in SN-related research [35]. Generally, prediction results are generated by training, testing, learning, and classifying methods based on a set of characteristics as input. Link prediction usually focuses on invisible edges [31] or new links that may emerge in the future. Popularity prediction, on the other hand, focuses on information popularity and its increasing rate among users.

In the homogenous environment of SNs, prediction commonly involves similarity calculation, node attributes, influential users, and intelligent algorithms. Non-intelligent approaches rely on two ideas: similarity [30] and key users [49], which are crucial factors in promoting and influencing interactions and diffusion in SNs. However, these approaches have obvious disadvantages when dealing with large SNs containing various information, such as the inability to handle explosive computations and complications, and limited effectiveness. Although intelligent algorithms, such as evolutionary algorithms, have partially alleviated these weaknesses to some extent [12], [26], there is still room for improvement.

Advancements in prediction in SNs have been brought about by combining machine learning methods with general prediction algorithms, particularly in the application of graph-based learning (GL) methods [41]. Traditional GL methods propose an approximating graph convolutional network for quick feature embedding or characteristic extraction. A supervised or semi-supervised method is then used to generate an advanced and appropriate prediction model. However, adjusting hyper-parameters and dealing with over-fitting risk in the learning process often leads to unsatisfactory predictions. While existing literature has made progress in relative theory and applications for various prediction problems [18], [19], little effort has been made to predict SNs under positive-negative diffusion (PND), which is a significant and common phenomenon in practical social activities.

Both positive and negative influences are dynamically diffused among a user's neighborhood, making it difficult to build a model in social networks and challenging to capture the features of the complex diffusion through the existed methods. This paper proposes an original MAS-based attention mechanism to address the prediction problem of product popularity under PND in SNs. The main contributions are listed as follows.

1) A novel PND model, which includes three attitudes of humans towards a new product, is proposed to simulate product popularity among an SN.
2) An MAS-based Attention Mechanism (MASAM) is proposed to extract precise characteristic representations of every user in the SN, forming the foundation of diffusions and predictions.

This work was supported in part by the National Key Research and Development Program of China under Grant 2021ZD40303; in part by the National Natural Science Foundation of China under Grant 62206091; in part by the Creative Research Groups Program of the National Natural Science Foundation of China under Grant 62321003; and in part by the Natural Science Foundation of Hunan Province of China under Grant 2023JJ40166.
(Corresponding author: Kenli Li.)

Mincan Li and Kenli Li are with the College of Computer Science and Electronic Engineering, Hunan University, Changsha, Hunan 410082, China; and also with the National Supercomputing Center in Changsha, Changsha, Hunan 410082, China. (emails: limc@hnu.edu.cn; lkl@hnu.edu.cn.)

Zidong Wang is with the Department of Computer Science, Brunel University London, Uxbridge, Middlesex, UB8 3PH, United Kingdom. (email: Zidong.Wang@brunel.ac.uk)

Xiangke Liao is with the Collaborative Innovation Center of High Performance Computing, National University of Defense Technology, Changsha 410073, China. (e-mail: xkliao@nudt.edu.cn).

Citation information: DOI: 10.1109/TNNLS.2023.3330100, IEEE Transactions on Neural Networks and Learning Systems

3) Based on the representations extracted by MASAM, MAS carries out the simulation of human interactions on the product, and the popularity predictions are reflected by the final situations of agents.

The rest of this paper is organized as follows. Section II discusses the related work on the prediction problem and advanced technologies on SN. Section III illustrates the prediction problem of product popularity in SN and relative preliminaries. Proposed attention mechanism, positive-negative diffusion model, and agent interaction rules for predicting product popularity in SN are provided in Section IV. Section V presents experiment results, data and parameter analysis, and comparisons of diverse baseline approaches. Finally, conclusions and future research topics are drawn in Section VI.

II. Related Work

In most existing literature, prediction problems in SNs rely primarily on similarity-based methods. In link prediction problems, new links are more likely to be established by similar user/node pairs than non-analogous pairs [8]. Additionally, the spread of information or influence is quicker among similar users as compared to general users. Various similarity metrics have constantly emerged, which are roughly categorized into local similarities (e.g. Common Neighbors (CN) [44], Jaccards Coefficient (JC) [27], and Hub Promoted Index (HPI) [48]) and global similarities (e.g. path similarity [43] and random walk [15]). On one hand, local-similarity-based approaches reflect faster, more effective and higher parallelism compared to global-similarity-based ones. On the other hand, the calculation of global similarities takes into consideration extra topologies, which generates positive results especially when using the information of longer path of nodes.

Learning-based techniques, such as supervised and semi-supervised classification methods, are another powerful and efficient approach. Random forest [13], decision trees [1], and k-nearest neighbors [28] are recommended to support prediction based on classification. Moreover, profile features, location features, and social features extracted from social networks are proposed to predict links, attributes, preferences, and popularity. These approaches are often implemented using traditional neural networks, such as Convolutional Neural Networks (CNNs), Deep Q-learning Networks (DQNs), and Recurrent Neural Networks (RNNs). Recently, the development of Graph Neural Networks (GNNs) has accelerated the achievement of high prediction accuracy in SNs. Graph Convolutional Network (GCN) learns nodes' characteristic representation from their neighbors, extracting and handling graph structures and information [40] which include both ordered non-Euclidean data but also unordered data. Multi-layer GCN is suggested for learning and predicting node representation in repost relations in time-stamped SN. Graph Attention network (GAT) suggests different attention indexes according to neighbor's contributions, enhancing the accuracy of information cascade prediction and social recommendations [38]. In addition, a multi-head self-attention mechanism is suggested for the transformer model to extract long-dependencies and bi-directional relationships in sub-graphs [34].

The prediction of SN's popularity has vast potential applications including item forecasting, product promotion, topic/rumor detection, and user recommendation. A content optimization platform has been developed in [3] to analyze the popularity of content pieces on social media. An online advertising prediction model has been designed in [10] to predict future hot items, enabling quick user response. The success rate of marketing campaigns in SNs can be predicted and monitored, improving online marketing [14]. Search algorithms have now considered different levels of popularity among users, improving the determination of search result priorities [11]. Prediction technologies in SNs have direct applications in box office forecasting [33], rumor detection [2], and topic prediction [42]. Additionally, a model for predicting customer churn has been developed and applied to a large scale telecommunication dataset [36].

III. Problem Formulation

A. Preliminaries

Potential user (P): in the initial phase of new product's information dissemination on social networks, the primary user population consists of potential users who have not yet tried the product. However, product marketing or the influence of their neighbors can potentially transform each potential user into a regular or repelling user. It is worth noting that a potential user cannot influence others unless its potential state is transformed by other states.

Regular user (R): as the product gains momentum, regular users emerge and access it frequently at each time step.

Repelling user (L): over an extended period of product diffusion, repelling users may emerge due to bad experiences or disadvantages, and they can spread negative information, potentially converting potential users into repelling users.

Positive-negative diffusion (PND): based on the definitions of the above three types of users, a positive-negative diffusion model can be created on SNs. During the product diffusion process, these users interact through links and any positive or negative information diffused on the links is represented as a probability. Regular users promote the product by spreading positive information, while repelling users discourage its usage by spreading negative information. When a node becomes a *potential user* for the second time, it can influence its neighbors to become potential users. Positive and negative diffusion in this model are defined relative to the influence of a regular user, with potential users having a positive diffusion on repelling users, and regular users receiving positive diffusion from potential users but negative diffusion from repelling users.

Remark 1: A simple example of PND among MAS is illustrated in Fig. 1, where each user is denoted as an agent in MAS, which has three states a_0, a_1 and a_2, representing potential, regular and repelling states, respectively. The repelling user a_2 receives positive diffusion from potential user a_0 and negative diffusion from regular user a_1, which are marked as dark orange lines; the potential user a_0 receives positive diffusion from regular user a_1 and negative diffusion from repelling user a_2, which are marked as black lines; the regular user a_1 receives positive diffusion from potential user a_0 and negative diffusion from repelling user a_2, which are marked as green lines. In addition, there is no diffusion between two users in the same state.

B. Problem statement

User/Agent: in an MAS model, every user is an agent $a_.$ characterized by $\{s_{a_.}, p_{a_.}, \theta_{a_.}\}$ where $s_{a_.}$, which takes value on p (potential), r (regular) or l (repelling), represents the current situation of $a_.$. $p_{a_.} = \{pr_{a_.}, rl_{a_.}, lp_{a_.}, rp_{a_.}, lr_{a_.}, pl_{a_.}\}$ denotes the possibilities of corresponding state transforming, see Fig. 2. For instance, $pr_{a_.}$ denotes the possibility that agent a_i transforms from a potential user into a regular user. The values of lr and lp are typically much smaller than other possibilities in practical human activities, as it is rare and difficult for people to switch from a repelling to regular or potential state. The set of transforming thresholds for the three states is denoted as $\theta_{a_.} = \{\theta_1^{a_.}, \theta_2^{a_.}, \theta_3^{a_.}\}$.

We are now ready to state *the prediction problem of product popularity under PND* as follows: given a population of potential users linked by a fixed SN: $G = \{V, E\}$ (where V is node set that denotes all users and E is edge set among users), let us activate a set

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TNNLS.2023.3330100, IEEE Transactions on Neural Networks and Learning Systems

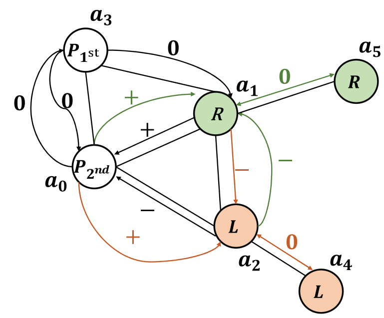

Fig. 1. An example for PND illustration.

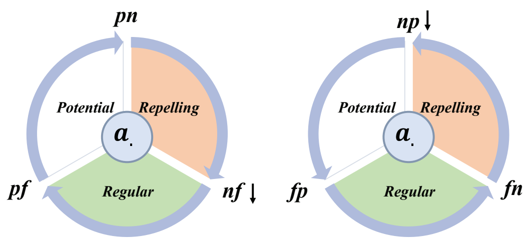

Fig. 2. The transforming of three states on every agent under corresponding possibility.

of original users OU as regular ones of the product. Then, we aim to calculate the prediction of product popularity (i.e. the percentage of regular users P_R) under PND after t_{max} time steps.

IV. Main Model and Approach

In this section, a PND model based on Potential-Regular-repeLling (PRL) users is proposed, along with interaction rules for MAS-based diffusion and an MASAM for product popularity predictions in SN.

A. PND model on PRL

The PRL diffusion model includes four steps: 1) initializing the population, 2) executing PND diffusion among users, 3) calculating the attenuation of information diffusion, and 4) analyzing the data statistics.

1) Population initialization: At the start of PRL diffusion, all users are considered potential, and the population $P = |V|$ is initialized by activating the nodes in OU. Specifically, the population is initialized as $|OU|$ regular users and $|V| - |OU|$ potential users. Repelling users are not set in this step, but they emerge after a period of time of information diffusion.

2) PND process among PRL: To calculate the influence sum received by each user during the PND process, the positive and negative diffusions on potential users are taken into account. For a potential user $a_.$ ($s_{a_.} = p$), the influence sum received can be calculated using the following equation:

$$sf_{a_.} = \sum_{a_i \in N_{a_.} \& s_{a_i} = r} pr_{a_i}\alpha_{.i} - \sum_{a_j \in N_{a_.} \& s_{a_j} = l} pl_{a_j}\alpha_{.j} \qquad (1)$$

where $N_{a_.}$ is the set of $a_.$'s neighbors, and $\alpha_{.i}$ and $\alpha_{.j}$ are the attention indexes that are from $a_.$ to a_i and a_j, respectively. The calculations of $\alpha_{..}$ will be introduced in Section IV-C.

If $sf_{a_.} = 0$, the agent will maintain its current state. Otherwise, by obtaining the value of $sf_{a_.}$, the transforming possibilities of $a_.$ are updated by:

$$\begin{cases} pr'_{a_.} = pr_{a_.} + sf_{a_.}, & \text{if} \quad sf_{a_.} > 0 \\ pl'_{a_.} = pl_{a_.} - sf_{a_.}, & \text{if} \quad sf_{a_.} < 0 \end{cases} \qquad (2)$$

Similar measures are applied to $a_.$, whose state is r (see (3)-(4)) or l (see (5)-(6)):

$$sf_{a_.} = \sum_{a_i \in N_{a_.} \& s_{a_i} = p} rp_{a_i}\alpha_{.i} - \sum_{a_j \in N_{a_.} \& s_{a_j} = l} rl_{a_j}\alpha_{.j} \qquad (3)$$

$$\begin{cases} rp'_{a_.} = rp_{a_.} + sf_{a_.}, & \text{if} \quad sf_{a_.} > 0 \\ rl'_{a_.} = rl_{a_.} - sf_{a_.}, & \text{if} \quad sf_{a_.} < 0 \end{cases} \qquad (4)$$

$$sf_{a_.} = \sum_{a_i \in N_{a_.} \& s_{a_i} = p} lp_{a_i}\alpha_{.i} - \sum_{a_j \in N_{a_.} \& s_{a_j} = l} lr_{a_j}\alpha_{.j} \qquad (5)$$

$$\begin{cases} lp'_{a_.} = lp_{a_.} + sf_{a_.}, & \text{if} \quad sf_{a_.} > 0 \\ lr'_{a_.} = lr'_{a_.} - sf_{a_.}, & \text{if} \quad sf_{a_.} < 0 \end{cases} \qquad (6)$$

3) State transforming: After obtaining the influence sum $sf_{a_.}$, we can calculate $a_.$'s transforming possibilities according to the following rules (7)-(9):

$$\text{If } s_{a_.} = p, \; s'_{a_.} = \begin{cases} r, & pl_{a_.} < pr_{a_.} \& \; pr_{a_.} > \theta_1^{a_.} \\ l, & pl_{a_.} > pr_{a_.} \& \; pl_{a_.} > \theta_2^{a_.} \\ p, & sf_{a_.} = 0. \end{cases} \qquad (7)$$

$$\text{If } s_{a_.} = r, \; s'_{a_.} = \begin{cases} r, & sf_{a_.} = 0 \\ l, & rp_{a_.} < rl_{a_.} \& \; rl_{a_.} > \theta_2^{a_.} \\ p, & rp_{a_.} > rl_{a_.} \& \; rp_{a_.} > \theta_3^{a_.}. \end{cases} \qquad (8)$$

$$\text{If } s_{a_.} = l, \; s'_{a_.} = \begin{cases} r, & lr_{a_.} > lp_{a_.} \& \; lr_{a_.} > \theta_1^{a_.} \\ l, & sf_{a_.} = 0 \\ p, & lr_{a_.} < lp_{a_.} \& \; lp_{a_.} > \theta_3^{a_.}. \end{cases} \qquad (9)$$

where $\theta_1^{a_.}, \theta_2^{a_.}$ and $\theta_3^{a_.}$ indicate the thresholds of $a_.$ on regular, potential and repelling states, respectively. $s'_{a_.}$ is the new state after transforming.

4) Diffusion attenuation: If a user remains in the same state for an extended period, the diffusion on that state decreases due to attenuation. The attenuation function shown in (10) describes this phenomenon. The time threshold for attenuation is set at 30 time steps.

$$\begin{cases} rp'_{a_.} = rp_{a_.}.exp(-t) \; \& \; lp'_{a_.} = lp_{a_.}.exp(-t), & s_{a_.} = p \\ lr'_{a_.} = lr_{a_.}.exp(-t) \; \& \; pr'_{a_.} = pr_{a_.}.exp(-t), & s_{a_.} = r \\ pl'_{a_.} = pl_{a_.}.exp(-t) \; \& \; rl'_{a_.} = rl_{a_.}.exp(-t), & s_{a_.} = l. \end{cases} \qquad (10)$$

where t is the time step and $t > 30$.

B. Main MAS model and interaction rules under PRL

Algorithm 1 presents the main MAS model [7], [16], [17]. In line 1, the initialization of MAS is executed by setting the social topology, attributes, and activating regular users. Then, the MASAM algorithm is invoked to extract features, as shown in line 2. At every time step, the PND model under PRL is executed by MAS through agent links in line 4. After diffusion, the current agent states are recorded in A'_{state}, and an attenuation function is applied to agents with long-lasting states. In the initialization step, there are no repelling users. Therefore, a trigger for the emergence of repelling users is designed in lines 7-9 by randomly converting regular users into repelling users at the 30^{th} time step. The states of all agents at the end of every time step are recorded in line 10, and the number of agents in each state is output by Algorithm 1.

The interactions among agents in the PND model are detailed in Algorithm 2. Each agent undergoes individual state judgment, confirmation, and diffusion attenuation in lines 3-5, followed by interactions with neighbors in lines 6-9. These individual managements

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TNNLS.2023.3330100, IEEE Transactions on Neural Networks and Learning Systems

Algorithm 1 Main MAS prediction algorithm

Input: A social network $G = (V, E)$, agents' attributes set P and Θ, activating set OU and time-step maximum t_{max}.

1: Initialize MAS: $A = V$, $A \leftarrow P, E\&\Theta$, $A_{state} \leftarrow OU$;
2: FeatureExtraction: MASAM();
3: **for** t=1 to $t = t_{max}$ **do**
4: PND interactions: AgentIntercation(PRL$\leftarrow$ PND);
5: A'_{state} $\leftarrow$transforming(A_{state}, Θ);
6: Attenuation(a.StateTimes()>30);
7: **if** t==30 **then**
8: RepellingRandomly($A_{t=30}$);
9: **end if**
10: A.record();
11: **end for**
12: Statistic: Account-r(); Account-p(); Account-l().

Output: $Account_r, Account_p, Account_l$.

are necessary for agent interactions. After the PND interactions, each agent updates its possibilities and undergoes state transformation if necessary using (10), see lines 10-11.

Algorithm 2 Agent interaction algorithm

1: **for** every $a. \in A$ **do**
2: **for** t=1 to $t = t_{max}$ **do**
3: IsInfluencer(if($sa. = p$));
4: $s_{a.}$ $\leftarrow$UpdateSate();
5: AttenuationIndex();
6: **for** every $a_i \in N_{a.}$ **do**
7: IsPN(a.,a_i); diffusion$(a_i, 2 - p(s_{a.}))$;
8: IsPN(a_i,a.); receiving$(a_i, 2 - p(s_{a.}))$;
9: **end for**
10: $p'_{a.} \leftarrow sf'_{a.} \leftarrow \sum receiving()$;
11: Transforming$(a., p'_{a.})$;
12: **end for**
13: **end for**

C. MAS-based attention mechanism (MASAM)

The most crucial procedure for popularity prediction under PND is determining the six possibilities included in $p_{a.}$. Therefore, $p_{a.}$ is considered a feature of the agent, and a **Feature Vector** is defined as $f_{a.}^{\top} = \{pr_a., rl_a., lp_a., rp_a., lr_a., pl_a.\}$. As users in SN are mainly influenced by neighbors, an MASAM is presented in this section for feature extraction and calculation of the attention index that supports the computations of influence sum in (1), (3), and (5).

The basic process of MASAM is illustrated in Fig. 3. The attention of each agent is focused on its neighbors such as a_0. A new feature, $f'a_0$, is generated by a_0 through the calculation of attention indexes α_{01}, α_{02}, and α_{03} using a shared matrix W. These attention indexes are depicted as orange lines in Fig. 3. Similar procedures are carried out for a_1, a_2 and a_3, which are represented by blue lines.

The basic process of MASAM is illustrated in Fig. 3. For each agent, its attention is paid on its neighbors such as a_0. A new feature f'_{a_0} is generated by a_0 through attention indexes α_{01}, α_{02} and α_{03} by calculating with a shared matrix W, which are marked as orange lines in Fig. 3. Similar measures are applied to a_1, a_2 and a_3, which are marked as blue lines.

The generation of attention indexes is divided into two steps: basic index computation and final index normalization. For each agent $a.$ and its neighbor a_i ($a_i \in N_{a.}$), the basic index of $a.$ is denoted as $\alpha^b_{.i}$, which can be calculated by:

$$\alpha^b_{.i} = \omega(Wf_{a.} \| Wf_{a_i}) \tag{11}$$

where $\|$ is the concatenation operator, ω is the weight vector for feature vectors and its dimension is 12. W is a global shared weight matrix that is of dimension 6×6 and can be trained. By obtaining the basic indexes of $a.$ to its neighbors, the final index (i.e. $a.$'s attention indexes $\alpha_{.i}$ on a_i) is calculated by:

$$\alpha_{.i} = \frac{exp(LeakyReLU(\alpha^b_{.i}))}{\sum_{a_j \in N_{a.}} exp(LeakyReLU(\alpha^b_{.j}))} \tag{12}$$

With the generation of every attention index on $a.$'s neighbors, $a.$'s new feature vector is generated by:

$$f'_{a.} = LeakyReLU\left(\sum_{a_i \in N_{a.}} \alpha_{.i} W f_{a_i}\right) \tag{13}$$

In Algorithm 3, the MASAM steps are provided. Basic indexes are obtained from neighbors and gathered in set AC, as shown in lines 3-6. Attention indexes are obtained and utilized for the calculation of new features, as depicted in lines 7-10. The attention indexes and new features will be recorded, as indicated in line 11.

Algorithm 3 MASAM Algorithm

Input: A social network $G = (V, E)$, Feature-weight vector ω.

1: **for** every $a. \in MAS$ **do**
2: $f_{a.}$=Initialize(a.);
3: **for** every $a_i \in N_{a.}$ **do**
4: $\alpha^b_{.i} \leftarrow \omega(Wf_{a.}, Wf_{a_i})$;
5: $AC_{a.} \leftarrow \alpha^b_{.i}$;
6: **end for**
7: **for** every $a_i \in N_{a.}$ **do**
8: $\alpha_{.i} \leftarrow$ AttentionCoefficient($AC_{a.}$);
9: **end for**
10: $f'_{a.} \leftarrow Feature(W, f_{a_i})(a_i \in N_{a.})$;
11: a..Records();
12: **end for**

In Fig. 4, the training process of W is depicted. The computation of new feature vectors by agents is performed using the shared W. Subsequently, the feature loss is calculated by labeled agents, which are represented by orange circles. The loss value $\sum_{p=1}^{6} Lf_{a_1}(p) In(f'_{a_1}(p))$ is computed by labeled agent a_1. Similar loss computations are carried out by a_1, a_3, and $a.$, respectively, and collected as the final loss for updating W. The loss function is presented in

$$L = -\sum_{k \in A_L} \sum_{p=1}^{6} Lf_{a_k}(p) In(f'_{a_k}(p)) \tag{14}$$

where A_L indicates the set of labeled agents, Lf_{a_k} denotes the labeled feature vector of a_k and $Lf_{a_k}(p)$ is the p^{th} value of Lf_{a_k}.

The loss calculation is utilized for the updating of W based on the label vector. The training steps for the weight matrix W are outlined in Algorithm 4 according to the gradient descent method.

V. Experiments and Analysis

In this section, the experiment environment, database, baseline methods, and evaluation metrics for the issue of popularity prediction are presented in Section V-A. Subsequently, the PND model designed under PRL and MASAM, applied on a fixed SN, is tested under various experimental settings in Section V-B. The performance analysis,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TNNLS.2023.3330100, IEEE Transactions on Neural Networks and Learning Systems

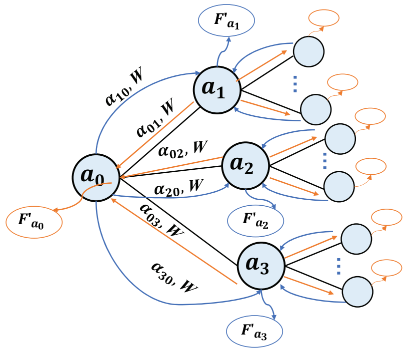

Fig. 3. Feature calculation in MASAM under SN.

Algorithm 4 Weight Matrix Learning Algorithm

```
for t=1 to t = t_max do
    if t == 1 then
        Initialize(W);
        Loss=0;
    else
        for a. ∈ MAS do
            a..MASAM();
            if IsLabeled(a.) then
                Loss=Loss+a..loss();
            end if
        end for
        W' ←Update(W,Loss);
    end if
end for
```

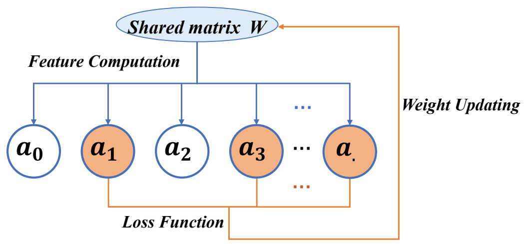

Fig. 4. Training process of W.

relative to baseline methods, is provided in Section V-C. Finally, a case study on the social campus network is disclosed in Section V-D. The proposed algorithms are implemented through programming in C++ and Python, and the algorithm execution is performed on CPU equipment.

A. Database, evaluation metrics and baseline methods

1) Database: The experiment database is collected from a new forum platform using a Python crawler. This dataset contains social relationships, accessing frequencies, and existence lengths of 2000 users over a span of 300 days. Specifically, the user relationship is indicated by their following queue in the data, where two users following each other signifies a link in this social network (SN). The accessing frequency is obtained by counting the user's login time on the forum within a day. The existence lengths of each user represent the day intervals of continuous access to the forum.

2) Evaluation metrics: The evaluation metrics include: i) percentages of regular users, ii) confusion matrix, and iii) precision.

i. Statistics about regular users' percentage

a) the maximum percentage of regular users among t_{max} time steps is

$$R_{max}\% = MaxNumber(R)/|Population| \tag{15}$$

b) the average percentage of regular users among t_{max} time steps is

$$R_{avg}\% = \frac{\sum_{t=1}^{t_{max}} Number(R,t)/|Population|}{t_{max}} \tag{16}$$

c) the average time of transforming to regular state for every user in the population is

$$R_{trans} = TotalNumber(R_{trans})/|Population| \tag{17}$$

ii. Confusion matrix By establishing a confusion matrix of PRL users, the classification precision of each type of user becomes evident.

iii. Precision The precision calculated in this paper is macroscopic, focusing solely on the percentage precision of the final popularity after diffusion:

$$Pre\% = 1 - |R_{Simulation} - R_{True}|/R_{True}\% \tag{18}$$

3) Baseline methods: The following three models are considered here.

Liner threshold (LS) model: In the LT model, the user's influence is accumulated from its neighborhood. Once the influence accumulation exceeds its threshold, the diffusion is accepted by the user. Similarly, in the experiments, the influence accumulation follows the PND diffusion. Furthermore, each user has a positive threshold for its regular state and a negative threshold for its repelling state. If the user's influence is zero, it becomes a potential user.

User attributes (UA) model: Users with similar attributes have the possibility of remaining in the same state for new product diffusion. In this method, the similarities of historical behaviors and existence lengths are calculated for users. Based on these similarities, the positive and negative attitudes towards the product are diffused among similar users.

Graph embedding (GE): Considering the process of random walk in GE, the walking probabilities are calculated for diffusion probabilities. The PND is then carried out among users and the popularity is predicted based on these probabilities.

B. Experiments on various settings for proposed models

To evaluate the proposed models and algorithm, the focus of this section is on the experiment results under various settings. Firstly, the performance of metric i, which is the percentages of regular users, is presented in Fig. 5.

In Fig. 5, the practical percentages of regular users are also displayed. It is evident that the prediction results deviate slightly from the real data. However, the percentages are controlled within a reasonable range for predicting popularity on R_{max} and R_{avg} with different population sizes. The average time taken for each user to transition to a regular state indicates the extent of attitude exchange among users. The longer the average time, the more pronounced the emotional fluctuations in user attitudes. Consequently, the product diffusion becomes more unstable and requires more time to reach a final equilibrium.

Secondly, confusion matrices are established in Table I for various population sizes. In each row, under the fixed population, the sum of percentages is 100%, and every value represents the percentage of predicted nodes for each category. It is worth noting that the True Positive Rates, such as among 300 users, are displayed as 55.6%, 75.4%, and 89.0%. The macroscopic prediction is the main focus of

Citation information: DOI: 10.1109/TNNLS.2023.3330100, IEEE Transactions on Neural Networks and Learning Systems

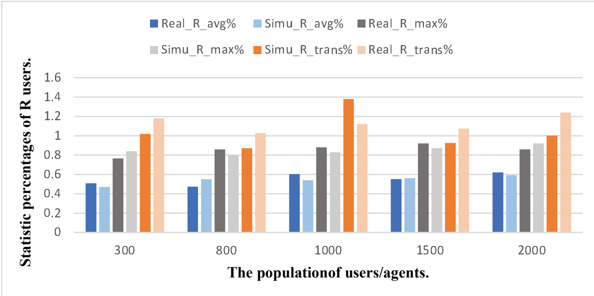

Fig. 5. The performance of percentages of regular users.

all proposed methods, and the results in the confusion matrices are microscopic, which may not be highly satisfactory but are acceptable. The changing process of PRL among 1000 users over 245 time steps is presented in Fig. 6.

TABLE I
CONFUSION MATRIX UNDER DIFFERENT POPULATION SIZES.

	Population size: 300			**Population size: 800**		
True \ Pred	P	R	L	P	R	L
P	55.6%	25.2%	19.2%	65.5%	14.7%	19.8%
R	17.2%	75.4%	7.4%	13.6%	78.3%	8.1%
L	8.7%	2.3%	89.0%	10.8%	6.4%	82.8%
	Population size: 1000			**Population size: 1500**		
True \ Pred	P	R	L	P	R	L
P	71.9%	20.4%	7.7%	69.2%	15.5%	15.3%
R	13.5%	83.6%	2.9%	9.8%	87.7%	2.5%
L	1.1%	13.7%	85.2%	7.1%	4.1%	88.8%
	Population size: 2000					
True \ Pred	P	R	L			
P	58.5%	35.2%	6.3%			
R	9%	83.3%	7.7%			
L	14.9%	9%	76.1%			

Finally, the results of the last metric Pre are displayed in Fig. 7, represented by blue bars. It is satisfying to observe that the average precisions from 30 running times fall within the range of $(82 \backsim 92\%)$. The gray bars represent the final percentages of regular users in the true database, while the orange bars represent the predictions.

C. Comparison with baseline methods

The comparison of the LT model, AU method, GE, and MASAM is illustrated in Fig. 8, where deviations from the true database are calculated. It can be observed that MASAM generates the minimum deviation under every population size. The LT model exhibits the maximum deviation because the information diffusion is successful at every time step, leading to continuous accumulation of influence. Once a user becomes a potential user in the LT model, it becomes challenging to decrease their threshold to a negative value and transition to a repelling user. The second poorest performance is observed in the AU method, as the prediction becomes uncertain when similar users of one user have different states. The diffusion probabilities calculated by GE result in approximate but inferior outcomes compared to MASAM.

D. A case study on a social campus network

A case study is conducted on the popularity prediction of a mobile-phone application called "curriculum management" using the proposed methods. The simulation involves the promotion of the

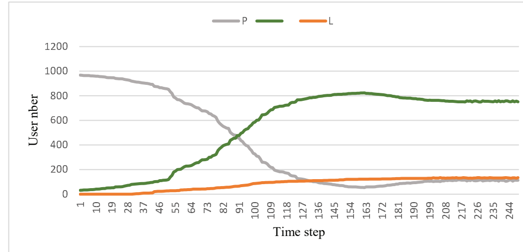

Fig. 6. The changing of three-state users in 1000 population.

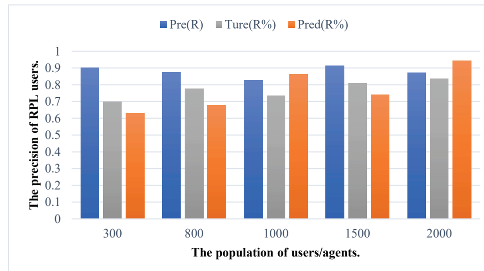

Fig. 7. The precision results on different populations.

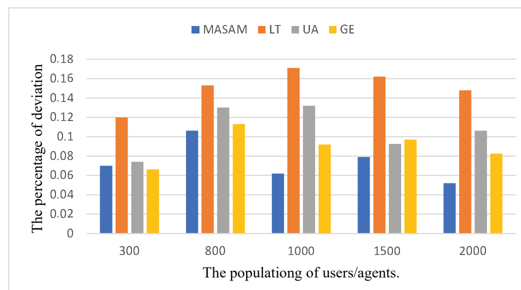

Fig. 8. Comparison of various methods.

"curriculum management" application through PRL-based PND and MASAM methods. The aim is to predict the popularity of the "curriculum management" application among all users of a social campus network. The initialization of the feature vector for each user is based on their historical behaviors on new applications, calculating the average probabilities of being a potential, regular, and repelling user for different products.

The first test is conducted on a randomly selected population of 500 individuals from the campus network. Then, 3% of users are randomly assigned as the original regular users, indicated by green nodes in Fig. 9(a). PND is performed using MASAM on the social campus network, and after 80 time steps, the results for the three types of users are displayed in Fig. 9(b). At the 80th time step, there are 22.8% repelling users and 53.6% regular users. Upon completion of 300 time steps, the final result of the popularity prediction is presented in Fig. 9(c), with a regular user percentage of 82.2%. The same test is repeated 50 times (with different original regular users in each run) to obtain the average popularity percentage of 76.7%, which is close to the actual value of 72.4%.

The second experiment involves 1000 users from the campus network. Considering the increase in population, the original user percentage is set at 8%. The results from one of the 50 running times are shown in Fig. 10. At the time step of 80, Fig. 10(b) displays 46.7% regular users and 43% repelling users. Comparing it to Fig. 9(b), it can be observed that the number of repelling users in the 1000 population is significantly higher, indicating a potential decrease

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/TNNLS.2023.3330100, IEEE Transactions on Neural Networks and Learning Systems

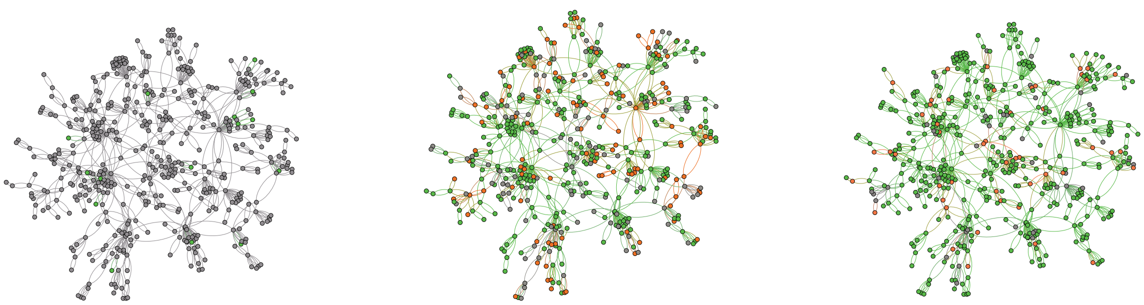

Fig. 9. The results of case study: 500 users.

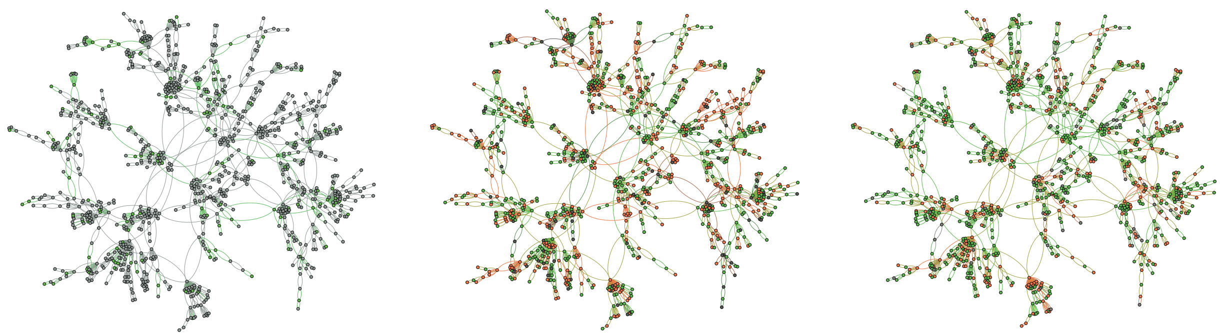

Fig. 10. The results of case study: 1000 users.

in product diffusion due to the increasing number of repelling users in this run. The final result is presented in Fig. 10(c), where the percentage of regular users is 60.8%. By running the algorithm 50 times, the average popularity is calculated to be 67.4%, while the actual proportion is 70.1%.

VI. Conclusion

This paper has addressed the prediction of product popularity on social networks (SN) using PND, employing MASAM and the PRL diffusion model. The three attitudes towards a product have been modeled as three types of users: potential users, regular users, and repelling users. The PRL diffusion model has been established based on the definition of PND to facilitate product spread on SN. The MAS model has been proposed for feature extraction through six possible transformations of the feature vector, considering the agent attributes. Through the distributed calculation of attention indexes and new feature vectors for agents, an optimal feature mapping model has been generated using a semi-supervised learning process. MAS has provided satisfactory predictions by utilizing the new feature as the primary feature for each agent. The experimental results and analysis have demonstrated the effective resolution of the product popularity prediction problem on SN using the proposed models and algorithms.

In conclusion, comprehensive experiments and a case study have revealed the effectiveness and efficiency of the prediction results. The proposed models and algorithms have exhibited significant potential for reference and extension. However, it is unfortunate that this study is limited by the lack of path prediction for information diffusion. Therefore, further work can be recommended to address path prediction and path similarity in information diffusion based on MAS, and it may combine with tenor model to improving computation in a high-dimensional environment [4], [9], [21], [23]–[25], [29], [37], [39], [45].

References

[1] A. Akavia, M. Leibovich, Y. S. Resheff, R. Ron, M. Shahar and M. Vald,“Privacy-preserving decision trees training and prediction,” *ACM Transactions on Privacy and Security*, vol. 25, no. 3, pp. 1–30, 2022.

[2] M. Al-Sarem, M. Al-Harby, F. Saeed and E. A. Hezzam,“Machine learning classifiers with pre-processing techniques for rumour detection on social media: an empirical study,” *International Journal of Cloud Computing*, vol. 11, no. 4, pp. 330–344, 2022.

[3] N. J. G. Amala and K. Kumar,“Content Popularity Prediction Methods-A Survey,” in *2018 3rd International Conference on Communication and Electronics Systems (ICCES)*, pp. 749–753, 2018.

[4] Y. Bi, Y. Yin, X. Liu, X. Nie, C. Zou and J. Du, “Real-time traffic flow statistics based on dual-granularity classification,” *International Journal of Network Dynamics and Intelligence*, vol. 2, no. 3, art. no. 100013, Sep. 2023.

[5] J. Cao, Y. Wang, Z. Bu, Y. Wang, H. Tao and G. Zhu, “Compactness preserving community computation via a network generative process,” *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 6, no. 5, pp. 1044–1056, 2022.

[6] J. Cao, Z. Bu, Y. Wang, H. Yang, J. Jiang and H. Li, “Detecting prosumer-community group in smart grids from the multiagent perspective,” *IEEE Transactions on Systems Man Cybernetics: Systems*, vol. 49, no. 8, pp. 1652–1664, 2019.

[7] Q. Cui, K. Liu, Z. Ji and W. Song, Sampling-data-based distributed optimisation of second-order multi-agent systems with PI strategy, *International Journal of Systems Science*, vol. 54, no. 6, pp. 1299–1312, 2023.

[8] S. Feng, G. Cong, A. Khan, X. Li, Y. Liu and Y. M. Chee,“Inf2vec: Latent representation model for social influence embedding,” in *2018 IEEE 34th International Conference on Data Engineering (ICDE)*, pp.941–952, 2018.

[9] S. Feng, X. Li, S. Zhang, Z. Jian, H. Duan and Z. Wang, “A review: state estimation based on hybrid models of Kalman filter and neural network,” *Systems Science & Control Engineering*, vol. 11, no. 1, art. no. 2173682, 2023.

[10] Z. Gharibshah and X. Zhu:,“User response prediction in online advertising,” *ACM Computing Surveys (CSUR)*, vol. 54, no. 3, pp. 1–43, 2021.

[11] M. A. Goncalves, J. M. Almeida, L. G. P. Santos, A. H. F. Laender and V. A. F. Almeida,“On popularity in the blogosphere,” *IEEE Internet Computing*, vol. 14, no. 3, pp. 42–49, 2010.

Citation information: DOI: 10.1109/TNNLS.2023.3330100, IEEE Transactions on Neural Networks and Learning Systems

[12] F. Gou and J. Wu:,“Triad link prediction method based on the evolutionary analysis with IoT in opportunistic social networks,” *Computer Communications*, vol. 181, pp.143–155, 2022.

[13] R. Haffar, D. Sánchez and J. Domingo-Ferrer,“Explaining predictions and attacks in federated learning via random forests,” *Applied Intelligence*, vol. 53, no. 1, pp. 169–185, 2023.

[14] M. Hajarian, M. A. Camilleri, P. Díaz, Paloma and I. Aedo,“A taxonomy of online marketing methods,” *Strategic corporate communication in the digital age*, pp. 235–250, 2021.

[15] B. Han,“Diffusion-Limited Aggregation Model on River Network Prediction Random Walk Simulation,” in *Proceedings of the 7th International Conference on Cyber Security and Information Engineering*, pp. 287–291, 2022.

[16] F. Han, J. Liu, J. Li, J. Song, M. Wang and Y. Zhang, “Consensus control for multi-rate multi-agent systems with fading measurements: the dynamic event-triggered case,” *Systems Science & Control Engineering*, vol. 11, no. 1, art. no. 2158959, 2023.

[17] G. He, M. Hu and Z. Shen, “Consensus of switched multi-agents system with cooperative and competitive relationship,” *Systems Science & Control Engineering*, vol. 11, no. 1, art. no. 2192008, 2023.

[18] D. Hristovski, A. Kastrin and T. C. Rindflesch,“Implementing semantics-based cross-domain collaboration recommendation in biomedicine with a graph database,” *DBKDA 2016*, vol. 104, 2016.

[19] N. M. A. Ibrahim and L. Chen,“An efficient algorithm for link prediction in temporal uncertain social networks,” *Information Sciences*, vol. 331, pp.120–136, 2016.

[20] M. Li, Z. Wang, Q.-L. Han, S. J. E. Taylor, K. Li, X. Liao and X. Liu, Influence maximization in multi-agent systems by a graph embedding method: dealing with probabilistically unstable links, *IEEE Transactions on Cybernetics*, vol. 53, no. 9, pp. 6004–6016, Sep. 2023.

[21] J. Liao, H.-K. Lam, S. Gulati and B. Hayee, “Improved computer-aided diagnosis system for nonerosive reflux disease using contrastive self-supervised learning with transfer learning,” *International Journal of Network Dynamics and Intelligence*, vol. 2, no. 3, art. no. 100010, Sep. 2023.

[22] Q. Liu, S. Zheng, Z. Wang, N. D. Alotaibi and F. E. Alsaadi, “Distributed estimation for multi-agent systems with relative measurements and quantized communication: A feedback quadratic framework,” *International Journal of Robust and Nonlinear Control*, vol. 33, no. 5, pp. 3164–3184, Jan. 2023.

[23] X. Luo, H. Wu, Z. Wang, J. Wang and D. Meng,“A Novel Approach to Large-Scale Dynamically Weighted Directed Network Representation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 12, pp.9756–9773, 2022.

[24] X. Luo, H. Wu and Z. Li,“A Novel Approach to Nonlinear Canonical Polyadic Decomposition on High-Dimensional Incomplete Tensors,” *IEEE Transactions on Knowledge and Data Engineering*, 2022.

[25] X. Luo, Y. Yuan, S. Chen, N. Zeng and Z. Wang,“Position-Transitional Particle Swarm Optimization-Incorporated Latent Factor Analysis,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 8, pp.3958–3970, 2020.

[26] M. K. Manshad, M. R. Meybodi and A. Salajegheh,“A new irregular cellular learning automata-based evolutionary computation for time series link prediction in social networks,” *Applied Intelligence*, vol. 51, pp.71–84, 2021.

[27] S. Niwattanakul, J. Singthongchai, E. Naenudorn and S. Wanapu,“Using of Jaccard coefficient for keywords similarity,” *Using of Jaccard coefficient for keywords similarity*, vol. 1, no. 6, pp.380–384, 2013.

[28] M. Przybyla-Kasperek, K. F. Marfo,“Neural network used for the fusion of predictions obtained by the K-nearest neighbors algorithm based on independent data sources,” *Entropy*, vol. 23, no. 12, pp. 1568, 2021.

[29] A. Rajagopal and S. Chitraganti, “State estimation and control for networked control systems in the presence of correlated packet drops,” *International Journal of Systems Science*, vol. 54, no. 11, pp. 2352–2365, 2023.

[30] M. Richardson and P. Domingos,“Mining knowledge-sharing sites for viral marketing,” in *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp.61–70, 2002.

[31] P. Songmuang, C. Sirisup and A. Suebsriwichai,“Missing Link Prediction Using Non-Overlapped Features and Multiple Sources of Social Networks,” *Information*, vol. 12, no. 5, pp.214, 2021.

[32] R. Ureía, G. Kou, Y. Dong, F. Chiclana and E. Herrera-Viedma, “A review on trust propagation and opinion dynamics in social networks and group decision making frameworks,” *Information Sciences*, vol. 478, pp. 461–475, 2019.

[33] B. Usero, V. Herníandez and C. Quintana:,“Social media mining for business intelligence analytics: an application for movie box office forecasting,” *Intelligent Computing: Proceedings of the 2021 Computing Conference*, vol. 1, pp. 981–999, 2022.

[34] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser and I. Polosukhin:,“Attention is all you need,” *Advances in neural information processing systems*, vol. 30, 2017.

[35] D. Vega-Oliveros, L. Zhao, A. Rocha and L. Berton,“Link prediction based on stochastic information diffusion,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 8, pp. 3522–3532, 2021.

[36] W. Verbeke, D. Martens, B. Baesens, “Social network analysis for customer churn prediction,” *Applied Soft Computing*, vol. 14, pp. 432–446, 2014.

[37] D. Wang, S. Shi, J. Lu, Z. Hu and J. Chen, Research on gas pipeline leakage model identification driven by digital twin, *Systems Science & Control Engineering*, vol. 11, no. 1, art. no. 2180687, 2023.

[38] M. Wang and K. Li:,“Predicting information diffusion cascades using graph attention networks,” in *Neural Information Processing: 27th International Conference, ICONIP 2020, Bangkok, Thailand, November 18–22, 2020, Proceedings, Part IV 27*, pp. 104–122, 2020.

[39] Y. Wang, H.-J. Liu and H.-L. Tan, “An overview of filtering for sampled-data systems under communication constraints,” *International Journal of Network Dynamics and Intelligence*, vol. 2, no. 3, art. no. 100011, Sep. 2023.

[40] X. Wu, W. Zeng and F. Lin,“GCNCPR-ACPs: a novel graph convolution network method for ACPs prediction,” *BMC bioinformatics*, vol. 23, no. 4, pp. 1–14, 2022.

[41] F. Xia, K. Sun, S. Yu, A. Aziz, L. Wan, S. Pan and H. Liu,“Graph learning: A survey,” *IEEE Transactions on Artificial Intelligence*, vol. 2, no. 2, pp.109–127, 2021.

[42] M. Xu, J. Du, Z. Xue, Z. Guan, F. Kou and L. Shi:,“A scientific research topic trend prediction model based on multi-LSTM and graph convolutional network,” *International Journal of Intelligent Systems*, vol. 37, no. 9, pp. 6331–6353, 2022.

[43] R. Yang, C. Yang, X. Peng and A. Rezaeipanah,“A novel similarity measure of link prediction in multi-layer social networks based on reliable paths,” *Concurrency and Computation: Practice and Experience*, vol. 34, no. 10, pp.e6829, 2022.

[44] L. Yao, L. Wang, L. Pan and K. Yao:,“Link prediction based on common-neighbors for dynamic social network,” *Procedia Computer Science*, vol. 83, pp.82–89, 2016.

[45] X. Yi, H. Yu, Z. Fang and L. Ma, “Probability-guaranteed state estimation for nonlinear delayed systems under mixed attacks,” *International Journal of Systems Science*, vol. 54, no. 9, pp. 2059–2071, 2023.

[46] Y. Yuan, G. Ma, C. Cheng, B. Zhou, H. Zhao, H.-T. Zhang and H. Ding, “A general end-to-end diagnosis framework for manufacturing systems,” *National Science Review*, vol. 7, no. 2, pp. 418–429, 2020.

[47] Y. Yuan, X. Tang, W. Zhou, W. Pan, X. Li, H.-T. Zhang, H. Ding and J. Goncalves, “Data driven discovery of cyber physical systems,” *Nature Communications*, vol. 10, no. 1, pp. 1–9, 2019.

[48] R. Zeng, Y. Ding and X. Xia,“Link prediction based on dynamic weighted social attribute network,” in *2016 International Conference on Machine Learning and Cybernetics (ICMLC)*, vol. 1, pp.183–188, 2016.

[49] K. Zhao,J. Yen, G. Greer, B. Qiu, P. Mitra and K. Portier,“Finding influential users of online health communities: a new metric based on sentiment influence,” *Journal of the American Medical Informatics Association*, vol. 21, no. e2, pp.e212–e218, 2014.