Who's Next: Rising Star Prediction via Diffusion of User Interest in Social Networks

Xuan Yang, Yang Yang, Jintao Su, Yifei Sun, Shen Fan, Zhongyao Wang, Jun Zhang, Jingmin Chen

Abstract—Finding items with potential to increase sales is of great importance in online market. In this paper, we propose to study this novel and practical problem: rising star prediction. We call these potential items *Rising Star*, which implies their ability to rise from low-turnover items to best-sellers in the future. Rising stars can be used to help with unfair recommendation in e-commerce platform, balance supply and demand to benefit the retailers and allocate marketing resources rationally. Although the study of rising star can bring great benefits, it also poses challenges to us. The sales trend of rising star fluctuates sharply in the short-term and exhibits more contingency caused by some external events (e.g., COVID-19 caused increasing purchase of the face mask) than other items, which cannot be solved by existing sales prediction methods.

To address above challenges, in this paper, we observe that the presence of rising stars is closely correlated with the early diffusion of user interest in social networks, which is validated in the case of Taocode (an intermediary that diffuses user interest in Taobao). Thus, we propose a novel framework, RiseNet, to incorporate the user interest diffusion process with the item dynamic features to effectively predict rising stars. Specifically, we adopt a coupled mechanism to capture the dynamic interplay between items and user interest, and a special designed GNN based framework to quantify user interest. Our experimental results on large-scale real-world datasets provided by Taobao demonstrate the effectiveness of our proposed framework.

Index Terms—Online market, Sales prediction, Graph Neural Networks, Dynamic graph

✦

1 INTRODUCTION

It's not about what it is, it's about what it can become.
—Dr. Seuss

In the online market, the item sales change drastically over time. Evaluating whether the item has the potential to increase sales in the future is of great importance in online market. In this paper, we propose to study this novel and practical problem: rising star prediction. We call these promising items *Rising Star* for their ability to rise from low-turnover items to best-sellers in the future. For instance, the face mask becomes a rising star at the beginning of 2020, whose sales soared exponentially when COVID-19 began its rapid spread. The sales of face masks on Taobao reached 80 million from Jan.20 to Jan.21 2020, which is more than one thousand times of that of the last week. Although rising stars account for a small proportion of total items and their unstable sales trend adds more difficulties to prediction, the great value brought by them cannot be ignored. We outline the importance of rising stars by discussing following common issues in e-commerce.

Alleviate unbalanced recommendation. The rapid development of recommender systems in e-commerce has significantly helped online-stores to increase conversion rates and sales by matching users with relevant items. However,

most recommender systems tend to favor a small set of popular items in their recommendations, and disfavor currently unpopular items, even when these items could be potentially preferred by a significant number of users [1]. The "Matthew Effect" thereby becomes increasingly evident: the head items become more and more popular, while long-tail items lacking wide and timely exposure [2][3]. This issue has caused negative influence on the balance and sustainability of the online market ecosystem,. In this case, retailers of those long-tail items will turn to other competitors for better opportunities. Precise identification of rising stars can be used to find out the promising items in long-tail and give them enough exposure, which will balance the recommender system and benefit the whole business environment.

Balance supply and demand. The balance of supply and demand is important for online retailers. Untimely adjustment of supply would greatly affect the sales and thus decrease retailers' benefits [4]. This issue is more significant for rising stars, as the demand for which usually fluctuates sharply. For example, popular and fashionable clothes of the season may have great potential to be a rising star. However, if the retailers cannot predict the upcoming increase in sales, users will lost interest as the waiting time for delivery increases [5]. On the other hand, if the new item lack the potential to be a rising star, the retailers can prevent overproduction in advance by correct prediction.

Allocate marketing resources. In online markets, many promotion campaigns are conducted to stimulate sales [6]. As the high promotion cost and limited marketing resources, it is essential to select the right item with high return benefits. Compared with current best-seller, which have low marginal benefits, and other long-tail items without

- Xuan Yang, Yang Yang and Jintao Su are with the College of Computer Science and Technology, Zhejiang University, China. E-mail: xuany, yangya, jtsu@zju.edu.cn. Yang Yang is the corresponding author.
- Yifei Sun is with the College of Aeronautics and Astronautics, Zhejiang University, China. E-mail: sunyf@zju.edu.cn.
- Zhongyao Wang, Shen Fan, Jun Zhang, Jingmin Chen are with the Alibaba Group, China. E-mail: zhongyao.wangzy, fanshen.fs@alibaba-inc.com, socialnetwork@gmail.com, jingmin.cjm@alibaba-inc.com.
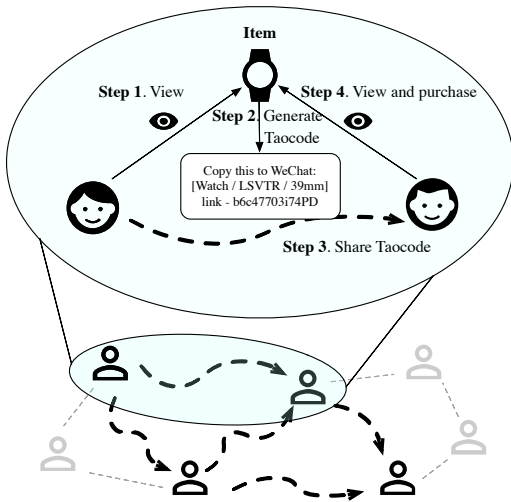
Fig. 1. Illustration of Taocode diffusion in social networks. A user can generate the Taocode message of an interested item and share it with specific friends. Afterwards, the receivers are able to further propagate the Taocode to their own friends, and thus creating diffusion. In the diffusion process, any receiver could open the Taocode, view and even purchase the item.

potential, rising stars would more likely to be the first choice of promotion campaign.

Up to this point, we have demonstrated the indispensable role played by rising star prediction in terms of fairness issue (from the perspective of the e-commerce platform), supply-demand issue (from the retailers' point of view) and marketing strategy. Still, finding the solution to the problem remains a great challenge. First, compared with other items, rising stars' sales trend tends to exhibit more suddenness caused by some *external events*. For instance, the aforementioned sudden rise in demand of face masks is caused by COVID-19. How might we capture such sudden sales changes without any further information about the external event? In addition, some rising stars tend to be new arrivals (e.g., iPhone 13) or items that just release from sales suspension; as a result, most of them lack long-term historical sales related data. Together with the consideration of external events influence, conventional sales prediction methods can not fit our rising star prediction problem, which mainly adopt time series analysis techniques and consider past long-term sales records.

To address above challenges, we propose and investigate the interplay between rising stars and the early diffusion of user interest in social networks. The idea is inspired by the following two aspects: 1) External events will be reflected by the diffusion of user interest [7][8]. For instance, COVID-19 led to the large-scale diffusion of user interest in the face mask. Therefore, even without the explicit information of external event, we could still perceive it by the diffusion of user interest. 2) Diffusion of user interest may prefigure future sales information [9][10]. Users that are activated in the diffusion process have a high probability of purchasing the diffused items in the future. This can by explained by principle of social proof: we tend to purchase items that others have purchased or recommended, especially people who are close or similar to us [11].

In order to study user interest diffusion, we conduct

a case study of Taocode. Taocode is used by more than hundreds of millions of users on Taobao to conveniently share items in various online social networks (e.g., Wechat[1] and Weibo[2]). Figure 1 demonstrates how Taocodes diffuse in social networks, a user generates the Taocode message of an interested item and shares with specific friends. Afterwards, the receivers are able to further share the Taocode to their own friends, and thus the diffusion formed. In the diffusion process, any receiver could open the Taocode, view the corresponding item, and even purchase it.

To conduct our study, we employ a real-world dataset, which contains 368 million records of Taocode diffusion, and relevant to 18 million items and 72 million users. In addition, we have 204 million purchase records. Our empirical observations show clear mutual interactions between Taocode diffusion and item purchase data: For instance, we find that diffusion scale (i.e., the number of users who receive the message) positively correlates with sales. Moreover, the sale volume of items propagated by "hub users" are more than five times of that of those haven't been propagated by "hub users". More importantly, we observe that rising stars tend to exhibit greater fluctuation in their diffusion statistics (e.g., diffusion scale), and they are more likely to be propagated by more "hub users" compared with other items. Also, we illustrate the importance of dynamic information in rising star prediction.

Despite of interesting insights provided by our observations, the question of how to jointly model the likelihood of an item to become a rising star and the dynamic Taocode diffusion into a unified framework is challenging. One straightforward method involves concatenating item historical data and statistical data of Taocode diffusion at each temporal point to capture the overall patterns. However, only using statistical data of Taocode diffusion will lose the rich information at user level. As Taocode diffusion data could naturally be regarded as dynamic graph, we propose a graph neural network (GNN) based model to quantify user interest. Unlike traditional GNN, we differentiate Taocode influence to users with consideration of both user item preference and Social Proof; and we also differentiate the roles played by different users in the diffusion graph to help graph information extraction. The item temporal data (such as sales) is essential in our task, the simple concatenation between the item and Taocode sources precludes capturing the significant mutual influences between them (Section 3). Here, we highlight our coupled mechanism in RiseNet, which combines information from both sides while modeling the interplay between them at each time step. As dynamic nature of our data, RiseNet further models the dynamic changes at each time step and adopts a multitask mechanism to better leverage the information. Experiment results on four Taobao real-world datasets demonstrate the effectiveness of our model RiseNet by achieving significant predicting score against existing alternatives.

Accordingly, the major contributions of our work can be summarized as follows:

• We propose to study the novel and practical problem: rising star prediction.

- Based on observational studies, we formulate a novel framework RiseNet to predict rising stars by leveraging user interest diffusion and item information simultaneously.
- Employing a large real-world dataset, we conduct extensive experiments to validate the superiority of our proposed model against several baselines.

2 PRELIMINARIES

2.1 Datasets

Our dataset comprises three sets of real-world data, all of which are provided by China's largest e-shopping platform, Taobao. We construct our dataset by sampling records during a time period of three continuous months in 2020.

Taocode diffusion data This dataset contains 368,888,906 records of Taocode diffusion among 72,594,418 users, involving 18,325,582 items. Taocode is a special form of sharing items among Taobao users which we have introduced in Section 1. We define a Taocode diffusion record for the shared item i as a 4-tuple $e = (i, u_s, u_r, t)$, where u_s denotes the Taocode sender, u_r denotes the Taocode receiver, and t is the timestamp when the record takes place.

Taobao purchase data This dataset covers 204,398,128 online purchase records on Taobao, involving 215,821,800 users and 116,867,875 items. Each purchase record contains the information of the buyer, the item, the turnover, and the purchase time.

Basic information data of Taobao items This dataset covers basic information of 31,093,128 items, such as name, price, and category in a three-level pattern. For example, basic information of an ice cream is recorded in the data as (Hagen-Dazs, 49, Food, Frozen food, Ice cream).

2.2 Definitions

In this section, we provide necessary related definitions. For notations, we use lower-case letters to denote scalars (e.g., k) or single object (e.g., i), and upper-case letters for sets (e.g., E).

Rising Star. *We define items that rise from low-turnover items to best sellers as rising stars.* To suit the demand of the application scenario of Taobao, we define regard an item that is not ranked within the top 3‰ of its high-level category in the first four weeks, but later rises to be ranked in the top 1‰ over the next two weeks as a rising star. Items are ranked in their own high-level categories. It is because there are differences between the average sale volume of different categories. For example, the average sales of items in food is more than three times higher than that in clothing.

Dynamic Taocode Diffusion Graph. For an item i, dynamic Taocode diffusion graph of item i is a sequence of Taocode diffusion graphs $G_i = \{G_i^0, G_i^1, G_i^2, ..., G_i^T\}$. We refer the graph constructed by item i's Taocode diffusion records at the time step t as the Taocode diffusion graph G_i^t. The Taocode diffusion graph $G_i^t = (V_i^t, E_i^t, M_i^t)$, where E_i^t denotes the set of Taocode diffusion edge, V_i^t denotes the set of users (nodes), and M_i^t denote the user feature matrix.

Rising Star Prediction. In light of above notations, we can define the rising star prediction addressed in this paper as follows: given a specific item i on the online platform, the goal is to estimate $P(y_i|G_i, X_i)$, which is the probability that item i will become a rising star ($y_i = 1$) or not ($y_i = 0$). X_i denotes item i's temporal features, which can be represented as a sequence $X_i = \{X_i^0, X_i^1, ..., X_i^T\}$.

3 EXPLORATORY ANALYSIS

In this section, we first study the interplay between user interest diffusion (Taocode diffusion) and rising star identification, and then observe the interplay between user interest diffusion and item features. Furthermore, we examine whether dynamic information is needed in rising star prediction.

3.1 Observing Rising Stars Through Taocode Diffusion

In order to study the relationship between rising star identification and Taocode diffusion graph, we respectively sample 2500 rising stars from clothing category during a period of specific six weeks in 2020, and 2500 normal items to conduct the observation experiments. Here, items that do not rise from low-turnover items to best sellers are defined as normal items. To comprehensively observe the correlation, we propose to investigate the Taocode diffusion graph from two different levels: graph level and node level.

Firstly, on the graph level, we aim to study whether the rising stars exhibit patterns that different from the normal items on diffusion graph. To examine this, we draw the distribution of the maximum graph scale difference between two consecutive weeks (MDSW for short) for the group of rising stars and the group of normal items during the specific six weeks.

$$MDSW = Max(s_i^m - s_i^n)_{1<=n<=5; m=n+1} \qquad (1)$$

where s_i^m denotes the sales volume of item i in week m.

In Figure 2(a), there is an obvious rightwards deviation of MDSW distribution of rising stars compared with that of normal items, which indicates the rising stars exhibit greater fluctuation in their diffusion graphs. This result shows that the Taocode diffusion graph can provide us with information to identify rising stars from the graph level.

Furthermore, we want to study whether the diffusion graph can provide information to identify the rising star from the node level. Here, we take the hub nodes as an entry point to examine this question. In the Taocode diffusion graph, hub nodes represent the most active users. In our experiments, we first quantify hub nodes as the users who have the top 1% rank number of Taocode sharing records in the Taocode diffusion data. We draw the distribution of the numbers of hub nodes that contained in each sampled rising stars' diffusion graph during the same specific six weeks 2020, and we draw the distribution of the normal items group for comparison. The distribution in Figure 2(b) illustrates that items whose diffusion graphs contain more hub nodes are more likely to be rising stars. From the subsequent case study, we find the reason behind this result. The hub nodes are the most active users on Taobao e-commerce platform, who have greater abilities to identify potential items depending on their abundant online shopping experience. Moreover, certain hub nodes (such as internet celebrities) can employ their significant online
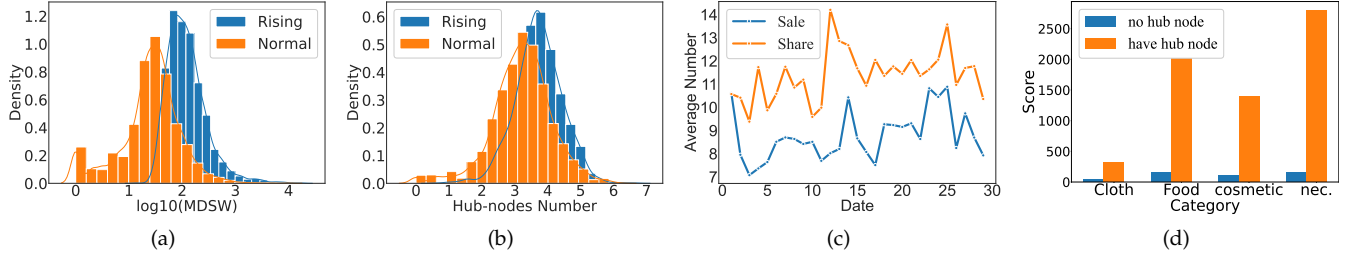
Fig. 2. Observation results of the Taocode diffusion network. (a) The distribution of maximum difference between the graph scales during two consecutive weeks (MDSW) of rising stars and normal items. (b) The diffusion graph hub node numbers' distribution of rising stars and normal items. (c) Interaction between item sales and diffusion graph scales. (d) Relationship between item sales and diffusion graph hub node numbers.

influence to recommend a particular item and thus create a rising star. To this end, we find that diffusion graph is also helpful from the node level in rising star identification.

Together, from above observations, we get the conclusion that the Taocode diffusion graph has close correlation with rising star identification from both graph level and node level.

3.2 User Interest Diffusion Vs. Item Features

As stated in Section 1, user interest diffusion is naturally related with item features. To further study this relationship, we also conduct observations on Taocode diffusion data from graph level and node level respectively.

First, from the graph level, we observe whether diffusion graph statistics (such as diffusion scale) correlate with item features (such as sales). We sample one million items from clothing category in our dataset and then examine their average sales and average diffusion graph scales that changes over one month in 2020. From the results shown in Figure 2(c), we can see that the item sales is directly proportional to the scale of Taocode diffusion graph, which validates the close correlation between item sales and user interest. This result is consistent with our recognition in real-world situations: item sales directly guide the interest of users due to the herd mentality [12] (users tend to pay more attention to items with high sales numbers because they think these items are more reliable); meanwhile, user interest diffusion greatly influences the item sales (for example, active sharing among users via Taocode can greatly promote item sales).

As we have observed that the scale of diffusion graph dynamically interact with item sales from the graph level, we furthermore study the relationship from the node level. Here, we also conduct observations on hub nodes as an example. Owing to the sales gap between different categories which we have discussed in Section 2, we respectively sample one million items from clothing, food, cosmetic and necessities categories, and then label hub nodes according to Taocode sharing ranks of each category. We divide the sampled items into two groups depending on whether or not they were recommended by hub nodes within the selected one month. We compare average sales of each group's containing items from four categories; results are shown in Figure 2(d). The items that have been recommended by hub-nodes are found to have far higher sales in all four categories. The reason is that hub nodes are typically influential users, such as, popular bloggers, who are more likely to diffuse selected items' Taocodes to their followers,

and professional buyers who actively send Taocodes to their customers. Users are not only more likely to receive Taocodes from these influential users, but also more willing to take their advice and then purchase the recommended items online. From the node level, we can see that Taocode diffusion graph also has its correlation with item sales.

From both the graph level and node level, we find close relationship between item features and user interest diffusion. Next, we want to explore whether the dynamic information is necessary in rising star prediction.
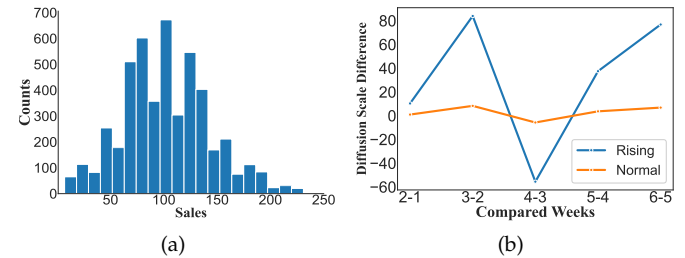
3.3 Importance of Dynamic Information

Fig. 3. Observation of the dynamic information. (a) The next week's sales diffusion of the items with the same sales the last week. (b) The average weekly diffusion scale difference over the last six weeks for rising star and normal items that have the same diffusion scale over the last week.

In online markets, the item information is naturally dynamic. It drastically changes with time, which includes sales, price, etc. What if we ignore this dynamic information and simply predict item sales with the information from last time step? To answer this question, from items in clothing category, we sample all items which have the same sales volume as average in a specific week. Then we observe the sales volume distribution of these items in the next week, shown in the Figure 3, which is not concentrated. This result suggests that items with same information in the last time step is much more likely to show divergent sales trend in the future. In other word, more historical information is necessary to make the prediction more accurate.

Like item features, user diffusion graph also evolves with time. Here, we investigate whether the dynamic information in diffusion graph can further help us to identify rising stars when the information from the last time step is not the only condition. We sample all items that have the same diffusion graph scale as average in a specific week, and divide them into two groups according to whether

they are rising stars in the next week. After that, we draw the average weekly diffusion scale difference (fluctuation) of each group with historical information in the last six weeks. As presented in Figure 3, although both groups have the same diffusion scale in the last week, rising star group shows more volatility in the diffusion scale difference in the last six weeks compared with the other group. This result shows that with the help of historical diffusion information, we can identify rising stars that cannot be identified with only condition of information from the last time step.

To this end, we illustrate the importance of utilizing the dynamic information in rising star prediction.

4 OUR APPROACH

In this section, with inspirations derived from Section 3, we propose our model, RiseNet, with the general idea to incorporate user interest diffusion with item information to predict rising stars. From the observation results in Section 3, we validate the following conclusions:

1) User diffusion graphs can provide great support for rising star prediction task from both graph-level and node-level.
2) There is a strong mutual interaction between the user interest diffusion and item features.
3) The dynamic information contained in item features and diffusion graph is helpful for rising star identification.

Inspired by above observation results, we propose the *RiseNet* model.

4.1 General description

Here, we elaborate the general framework of RiseNet. As illustrated in Figure 4, RiseNet aims to incorporate both the diffusion process of user interest and the dynamics of item features to effectively predict rising stars.

As the diffusion process could naturally be regarded as dynamic graphs, we propose a graph neural network based framework to fully capture the information in Taocode diffusion from both graph and node level. Unlike the traditional GNN models, we differentiate the roles of different users when we generate the graph representation and model each Taocode diffusion influence in the propagation process with consideration of both user item preferences and *social proof*. Meanwhile, considering the importance of item temporal information, we also model the item information from item temporal features (such as sales and price). Then we combine the item and user interest information to get a unified representation while capturing mutual interactions between them at each time step. Considering the dynamic nature of our data, we further build the temporal dependency of the unified representation at each time step, capturing the dynamic changes to help identify rising stars.

For details, the core components of RiseNet are elaborated in the remaining of this section.

4.2 Modeling User Interest Diffusion Graph

To quantify the dynamic user interest information in Taocode diffusion graph, here we introduce our designed graph neural network based framework. Naturally, GNNs can be applied to quantify the user interest due to their ability to model the non-linear correlation of the graph structure and user interactions in diffusion. Also, in the GNN user propagation process, the model could capture the collaborative filtering signal and thus strength user interest learning and improve predicting effectiveness [13].

More specifically, we introduce the overall pattern of our GNN module. In the GNN propagation process, as different Taocodes affect user interest to varying degrees, it is intuitive that different Taocodes (edges) should be assigned to different weights. DiffGate models the various Taocodes' influence to users by considering both user item preference and user relationship influence. After propagation, we then generate the Taocode graph representation with the learnt node embeddings. In the diffusion graph, a small part of users (hub nodes) often play decisive roles in the overall user interest change (Section 3.2). Our designed PickGate picks these hub nodes and conducts the readout function on the hub nodes and all nodes respectively. Then we get the final graph representation by concatenating them. PickGate could avoid the important information contained in hub nodes to be over-smoothed by other nodes, which helps to retain more user interest information in the graph representation. In the GNN training process, multitask mechanism is adopted on GNN module, which helps GNN to converge the user interest information from dynamic Taocode graphs.

The details of GNN module are introduced in the following:

DiffGate. DiffGate models the various influence of different Taocodes to users, which helps to better capture user interest in the propagation process.

Generally speaking, a Taocode's influence on user interest is mainly affected by the following two factors: 1) dynamic item information. When Taocodes share different items, the influence they bring to users is naturally different. A user tends to pay more attention to the Taocodes which share his interested categories' items. For example, compared to Taocodes that share rock music CDs, those share piano CDs will attract classical music lovers more easily. Moreover, even with the same item, the observation results in Section 3.2 show that the user interest will be greatly influenced by the item's current features (such as sales and price); 2) relationship between the Taocode sender and receiver. According to the principle of social proof [11], user interest will change to varying degrees according to different relationships between the Taocode sender and the receiver. People tend to accept the Taocode recommendation from their close friends. Also, some special types of relationships, such as teacher-student relationship, will greatly affect the user's interest towards the Taocode.

Simultaneously considering the above two factors of influence, Diffusion gate (DiffGate) models the probability that a Taocode will influence a user's interest to help the GNN capture the user interest information in the propagation process. More specifically, we first calculate two weights respectively: item attention for modeling dynamic item information's influence; and relationship attention for modeling user relationship' influence. And then we unify these two weights to get the final user interest influence possibility of a Taocode. For a Taocode, which shares item i between the user u and the user v, we define its item
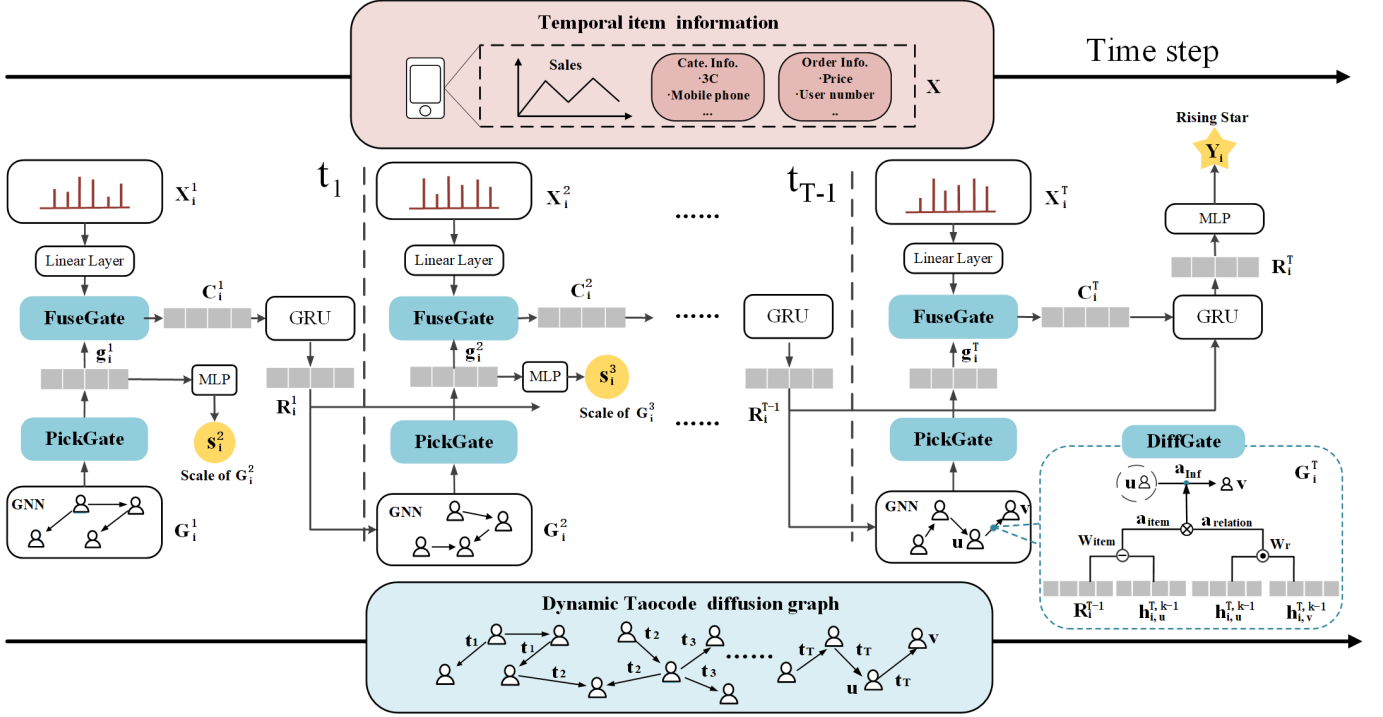
Fig. 4. The structure of RiseNet. Given dynamic Taocode diffusion graph and item temporal information, RiseNet captures the interplay between them through a coupled mechanism: a DiffGate (model various influence of Taocodes on user interest), a FuseGate (unify item and use interest information). A PickGate in GNN is to distinguish the roles of different nodes when generate the graph representation with node embeddings. A multitask framework is used in training to better leverage user interest in the dynamic diffusion graphs. The GRU cells are responsible for modeling temporal dependency.

attention for the user u as following,

$$a_{u,v,i^{item}} = W_{item} \cdot [h_{i,u}^{t,k-1} - I_i^{t-1}] \quad (2)$$

where $h_{i,u}^{t,k-1}$ represents the user u's hidden embedding in the G_i^t diffusion graph and layer $k-1$ of the GNN model. I_i^{t-1} is the representation of item i, which is generated at the last time step; this will be introduced in Section 4.3. W_{item} is a learnable vector to project the embedding to the influence probability. Note that we don't directly concatenate the item and user embeddings to do the projection. The reason is such attention could only get the global importance, as discussed in [14], and the ranking of the item attention scores is unconditioned on the query user, which is inconsistent with our intention to capture the various user interest towards Taocodes of different items.

For the relationship attention of a Taocode diffusion, we define it as:

$$a_{u,v,i^r} = W_r \cdot [h_{i,u}^{t,k-1} \odot h_{i,v}^{t,k-1}] \quad (3)$$

where W_r is a also learnable projection vector.

With the calculated item attention and relationship attention, we combine these two attentions and calculate the Taocode influence probability $a_{u,v,i}$ for user u as:

$$a_{u,v,i} = DiffGate(h_{i,u}^{t,k-1}, h_{i,v}^{t,k-1}, I_i^{t-1}) \quad (4)$$

$$DiffGate(h_{i,u}^{t,k-1}, h_{i,v}^{t,k-1}, I_i^{t-1}) = a_{u,v,i^{item}} \cdot a_{u,v,i^r} \quad (5)$$

With the calculated Taocode influence probability $a_{u,v,i}$, for each user u in a given user diffusion graph G_i^t, we first conduct the aggregation in each GNN layer:

$$h_{i,u}^{t,k} = h_{i,u}^{t,k-1} + \sum_{v \in N(u)} a_{u,v} \cdot h_{i,v}^{t,k-1} \quad (6)$$

Here, $N(u)$ denotes the set of neighbor nodes of u. The initial user representation is from M_i^t, the user feature matrix.

PickGate. After a fixed number L of layers of GNN propagation, We want to generate the Taocode graph representation with the learnt node embeddings. PickGate is designed here to retain more important information in the final graph representation to help rising star prediction. As discussed in our observations, there are some important users (named hub nodes) in the diffusion graph. They may be the internet celebrities or professional buyers, even the seller of the item. These hub nodes share a large amount of Taocodes and take the decisive role in the evolving of diffusion graph. Thus, instead of directly applying a readout function on all user embeddings to get the graph representation, PickGate picks these hub nodes in advance. Then we respectively apply the readout function to their embeddings and all nodes' embeddings respectively. And we get the final graph representation by concatenating them. The PickGate could avoid the important information that contained in hub nodes to be over-smoothed by other users. In the PickGate, we adapt a learnable vector to project the representation of each user to get their importance score [15]. According to the importance

score, we pick out top K% users,

$$idx = PickGate(\{h_{i,v}^L, v \in V\}, \lceil KS \rceil) \quad (7)$$

where S is the number of nodes.

Then we get entire diffusion graph representation g_i^t,

$$
\begin{aligned}
g_i^t =& READOUT(\{h_{i,v}^L, v \in V\}) || \\
& READOUT(\{h_{i,v}^L, v \in V(idx)\})
\end{aligned} \quad (8)
$$

Here, READOUT(*) is a summation function. We prefer the summation function to the average function, as we want to retain more global structural information.

Multitask mechanism. Compared with the static graph, the structure of dynamic graph changes drastically over time. In the GNN training process, the drastic changes on the dynamic diffusion graph make the amount of information provided by a single supervisory task insufficient for our task. What makes the task more challenging is that our target label is unbalanced due to the scarcity of rising stars. A designed multi-task mechanism helps the GNN model to capture the user interest more accurately while solving the problem of insufficient supervisory information.

Specifically, in addition to the global supervision task (rising star prediction) on dynamic graph at all time steps, we add a local supervision task (diffusion scale prediction) at each time step to help the GNN to better learn the user interest information. The diffusion scale prediction task can help the model to recognize the rising stars from the following aspects: 1) From the observation results in Section 3.1, the scale information of the diffusion graph can be used to distinguish rising stars. 2) The task can help the GNN to understand the user sharing and forwarding behaviors, which is helpful for the model to understand the user interest towards items.

More specifically, in each time step $t \in 1, ..., T-1$, we predict the scale s_i^{t+1} of the next time step's diffusion graph G_i^{t+1}. In this way, the scale prediction task helps us to find out the rising stars in a joint-training fashion with the target task. The predicted scale $\hat{s}_i^{t+1}$ is obtained via a multi layer perceptron:

$$\hat{s}_i^{t+1} = W_{s2} \cdot ReLU(W_{s1} \cdot g_i^t + b_{s1}) + b_{s2} \quad (9)$$

$Loss_0$ is defined as a mean relative square error (MRSE) loss, which is a common evaluation metric in scale prediction task:

$$Loss_0 = \frac{1}{N} \sum_{t=1}^{T-1} \sum_{n=1}^{N} (\frac{\hat{s}_i^{t+1} - s_i^{t+1}}{s_i^{t+1}})^2 \quad (10)$$

Here, N is the total number of dynamic diffusion graphs (batch size), while T is the total number of time steps.

4.3 Coupling Dynamic User Interest with Temporal Item Information

In addition to the user interest information contained in Taocode diffusion graph, the item's historical features are naturally important for the rising star prediction task. Thus, we aim to combine the information from both parties to predict the rising star accurately. As we discussed in Section 3, the item information and user diffusion information influence each other dynamically. Capturing this interplay is the key to enhance our model's ability in rising star prediction.

Therefore, we apply a couple mechanism to combine user interest and item information while capturing this sophisticated correlation between them. This mechanism comprises two main parts: *DiffGate* and *FuseGate*. In the coupled mechanism, firstly in each time step, the FuseGate combines the user interest and item temporal features' information while also capturing their mutual dependency to generate the item representation. And in the next time step, the DiffGate incorporates the generated item representation to calculate the item attention which helps to capture the user interest. Then the captured user interest information will be used as the input into the FuseGate at the next time step. In this way, a dynamic process of mutual influence is formed by our couple mechanism.

FuseGate. Once we have obtained the graph representation g_i^t generated by GNN at time step t, the next challenge we face is that of how to unify the graph information with the input of the item temporal features X_i^t. A fuse gate (FuseGate) combines the information from these two parties while also capturing their mutual dependency. In the fuse gate, we first use a learnable vector to get the item temporal representation from the initial features of the item. Then, the fuse gate models the mutual dependency of user interest and item information through a multi layer perceptron (MLP). After the MLP, the calculated vector is passed through a soft-max function, and then multiplied the concatenated graph representation and item representation to yield the representation of the combined information C_i^t:

$$C_i^t = FuseGate(g_i^t, X_i^t) \quad (11)$$

$$
\begin{aligned}
FuseGate =& SoftMax(MLP(g_i^t || W \cdot (X_i^t))) \\
& * (g_i^t || W \cdot (X_i^t))
\end{aligned} \quad (12)
$$

C_i^t denotes the unified representation of item i at time step t. which is then used by the dynamic model presented in Section 4.4 to generate item current presentation I_i^t. Item current presentation I_i^t will be used in the next time step's DiffGate. g_i^t is the generated graph representation with the help of DiffGate.

4.4 Capturing the Temporal Dependency

Our data is naturally dynamic. After exacting and unifying the information at each time step, and considering the importance of dynamic information (discussed in Section 3.3), a time series model allows us to flexibly utilize the temporal unified information C_i^t and build the temporal dependency of dynamic data among the time steps. Here, we take a recurrent model with GRU cells [16] as an example; at time step t:

$$z^t = \sigma(W_z C^t + U_z I^{t-1}) \quad (13)$$

$$r^t = \sigma(W_r C^t + U_r I^{t-1}) \quad (14)$$

$$\hat{I}^t = tanh(W C^t + U(r^t \odot I^{t-1})) \quad (15)$$

$$I^t = (1 - z^t) \odot I^{t-1} + z^t \odot \hat{I}^t \quad (16)$$

where $\odot$ is an element-wise dot operation. We share the GRU parameters for all the items and all the time steps. The initial input I_i^0 is a vector of all ones. Use of the GRU cell allows a flexible combination of previous item

TABLE 1
Overview of experimental datasets
(R. s. rate represents Rising star rate; sale in. means sales increase multiple from week one to week six)

Dataset	User	Taocode	Purchase	R. s. rate	sale in.
Cate-1	14,279,113	20,977,374	76,508,132	0.022	102.1
Cate-2	8,219,052	9,335,687	20,034,800	0.023	209.8
Cate-3	14,992,453	23,821,075	114,120,422	0.020	419.8
Cate-4	14,400,948	24,659,246	131,780,136	0.026	280.7

representations I_i^{t-1} and the newly arrived unified input C_i^t at time step t. GRU adopts two gates: the update gate z and the forget gate r. The forget gate r controls how much of the previous item representations I_i^{t-1} should be forgotten, while the update gate z controls how the newly arrived concatenated input C_i^t and the processed previous item representations I_i^{t-1} are combined.

After T time steps, by means of the iterative interplay between GNN and GRU cells, we obtain our final representation I_i^T for an item i. Then, we transform the representation I_i^T via a multi layer perceptron and a softmax function to get the predicted result p_i,

$$\hat{y}_i = Softmax(ReLU(W_2 \cdot ReLU(W_1 \cdot I_i^T + b_1) + b_2)) \quad (17)$$

We define the rising star prediction loss as a cross-entropy loss:

$$Loss_1 = -[y_i log(\hat{y}_i) + (1 - y_i) log(1 - \hat{y}_i)] \quad (18)$$

where y_i is the true label of the rising star prediction task.

Note that $Loss_0$, which is introduced in Section 4.2, is only used to update the GNN's parameters in a joint-training fashion with $Loss_1$. The loss of the graph neural network is defined as:

$$Loss_g = Loss_0 + Loss_1 \quad (19)$$

In this way, the graph neural network are trained jointly by the prediction of rising star and the next time step's scale prediction task. The multi-task learning further improves the generalize ability of our model.

5 EXPERIMENTS

In this section, we perform experiments to comprehensively evaluate our proposed model RiseNet. We aim to answer the following questions:

- **Q1**: How does the overall performance of our model compared with other state-of-the-art methods in rising star prediction task?
- **Q2**: Does user interest in Taocode diffusion help us with predicting rising stars?
- **Q3**: Does each component of our model RiseNet (i.e., multi-task framework, PickGate, coupled mechanism, hierarchical dynamic structure) contribute to our model on the target task?
- **Q4**: Does the dynamic information contribute to our task?
- **Q5**: What is the effect of changing key hyperparameters such as the GNN layer size for RiseNet?

5.1 Experimental setup

Experiment Datasets. We construct four experiment datasets on the basis of different high-level categories to evaluate the effectiveness of our model. Due to data privacy issues, we obfuscated the exact name of each category in this paper. In the process of constructing each dataset, for the diffusion graph, we sample 10 thousand items that have more than one hundred Taocode sharing records from a time period of 12 weeks in 2020. Then, according to the sampled items, we create our four experiment datasets based on the Taocode diffusion data, Taobao purchase data and Taobao item basic information data (see details in Section 2) during the same time period of 12 weeks in 2020.

The overall statistics of the sampled datasets are presented in Table 1, with obfuscated name of each category. We extract the temporal features of the sampled items in each time step, each of which mainly includes price, sales and category information. Moreover, the extracted features of users, who are engaged in the diffusion process, consists of information about users' purchase and diffusion habits. The purchase information includes user consumption level, category distribution of purchase product and consumption frequency, which helps the model to predict the probability that users will buy the items in the future. The diffusion habits include users' willing to share or receive Taocodes, forwarding probability and relationship between user purchase and sharing behavior, which help the model to forecast the future diffusion network and item sales.

Comparison methods. We compare our proposed model with the following state-of-the-art methods according to various research disciplines.

Sales Prediction Methods: The sales prediction task is usually considered to be a time series prediction problem. It has developed over time from utilizing conventional regression models [17][18] to employing deep neural network [19][20]. Here, we considered the following state-of-the-art methods that performs well in the sales prediction domain. Note that, due to these models' inability to handle the graph data, we concatenate statistics of graph (e.g., scale, breadth) with item features as input of time series model at each time step:

- **DA-RNN[21]**: This is a non-linear autoregressor (AR) with attention mechanism in both encoder and decoder RNNs. Compared with A-RNN, DA-RNN assumes that the input features must be correlated over time, and uses an encoder attention to capture the dependencies.
- **LSTNet[20]**: LSTNet is a deep learning framework (long- and short-term time series network), which is designed for multivariate time series prediction. This method combines a convolutional neural network and a recurrent-skip network to capture both short-term and long-term trending patterns of the time series.
- **MTGNN[22]**: MTGNN is the first to address the multivariate time series forecasting problem via a graph-based deep learning approach. This method exploits the inherent dependency relationships among multiple time series.

Graph Neural Network Models: To further evaluate our model, especially the ability to quantify user interest in diffusion graph, we consider to compare RiseNet with

TABLE 2
Rising star prediction results on four datasets

Method	Cate-1			Cate-2			Cate-3			Cate-4		
	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score
RiseNet	0.4072	0.6476	**0.5000**	0.2058	**0.4488**	**0.2822**	**0.1911**	0.3093	**0.2362**	**0.1640**	**0.4860**	**0.2453**
DA-RNN	0.3478	0.6038	0.4414	0.1582	0.2333	0.1886	0.1429	0.3438	0.2018	0.1171	0.3271	0.1724
LSTNet	0.2872	0.4297	0.3443	0.0765	0.3494	0.1225	0.0560	0.2110	0.0880	0.1000	0.4188	0.1614
MTGNN	**0.4080**	0.4881	0.4444	0.0762	0.2892	0.1206	0.0743	0.3846	0.1245	0.0879	0.4103	0.1448
GCN	0.2447	0.6571	0.3566	**0.2190**	0.2362	0.2273	0.1041	**0.3918**	0.1645	0.1134	0.3402	0.1701
GIN	0.2578	0.6286	0.3657	0.1638	0.2992	0.2117	0.1220	0.3196	0.1766	0.0989	0.4845	0.1643
DCRN	0.3857	0.5143	0.4408	0.1498	0.3858	0.2159	0.1176	0.3093	0.1705	0.1200	0.3093	0.1729
EvolveGCN	0.2537	**0.6571**	0.3661	0.1695	0.3937	0.2369	0.1352	0.3402	0.1935	0.0784	0.3814	0.1301

following two kinds of Graph Neural Network: 1) Static GNN model. For these methods, we first concatenate the graph representation learnt from GNN at every time step, and then concatenate it with the item features. We pass final representation through an MLP to conduct the rising star prediction. 2) Dynamic model. the implementation details are introduced below. The considered state-of-the-art methods are listed in the following:

- **GCN[23]**: From a spatial-based perspective, GCN can be considered as averagely aggregating the feature information from a node's neighborhood.
- **GIN[24]**: GIN adjusts the weight of the central node by means of a learnable parameter in the aggregation process in order to distinguish the graph structure information.
- **DCRN[25]**: The DCRN model combines convolutional neural networks (CNN) on graphs to identify spatial structures and recurrent neural networks (RNN) to find dynamic patterns. We concatenate the graph representation that generated by CNN with item temporal data, and then take the concatenated vector as the input of RNN at each time step.
- **EvolveGCN[26]**: It adapts the graph convolutional network (GCN) model along the temporal dimension without resorting to node embeddings. The EvolveGCN captures the dynamism of the graph sequence through using an RNN to evolve the GCN parameters. We concatenate final learnt representation of the dynamic graph and item temporal features. Then the concatenated representation is passed through a MLP to predict rising star.

Evaluation Protocols. We utilize three evaluation metrics to evaluate the performance of our model: Precision (how many positives were predicted correctly), Recall (the ability of the model to find all the positive samples), F1-score (calculates a mean of precision and recall). Considering the unbalanced nature of our labels, we pay more attention to F1-score.

Implementation details. Our experiment datasets contain datas that span 12 weeks in total. We use the first 8 weeks for training, weeks 5 to 10 for validating and weeks 7 to 12 for testing, so as to prevent label leakage. The utilized time steps' length for our experiment is set as four. The graph neural network layer number is set as two. We use the Adam optimizer with learning rate 0.001 to train the models. All the experiments are conducted on a 8-core CPU with 32 GB memory and Tesla V100 GPUs machine. Moreover, the algorithms are implemented in Pytorch [27] and Pytorch Geometric [28].

5.2 Overall Performance

In Table 2, we present the overall performance of comparison methods and our proposed model RiseNet on four real-world datasets in terms of Recall, Precision and F1-score. Based on these experimental results, we can answer the first research question:

1) Our model significantly outperforms other sales prediction domain methods in F1-score (improves 68.7% in F1-score on average). This result highlights the strength of our proposed RiseNet approach, which jointly captures dynamic user interest and item features. The State-of-the-art sales prediction methods such as LSTNet and MTGNN cannot handle our target task, which can be explained by the following two reasons. First, the development process from low-turnover items to rising stars lacks periodicity, which is the key information that comparison methods such as LSTNet want to capture. Secondly, as we mentioned in Section 1, items become rising stars occurs during a short time period that exhibits no long-term dependence. Therefore, MTGNN is not suitable in this case and do not perform well due to its over-fitting on long-term dependency.

2) Compared with the state-of-the-art GNN models, our model also achieves better performance in F1-score (improves 37.3% in F1-score on average). This demonstrates the superiority of our designed graph neural network model. For static models like GCN and GIN, their propagation methods lack the ability to model the heterogeneous influence of different Taocodes to users, which is a key to model the user interest towards items in diffusion. Also, they can not distinguish the roles of different users to retain key information when pooling node representations to graph representation. Besides, they do not capture the dynamic pattern of the evolving Taocode diffusion graphs. For the dynamic GNN models (such as DCRN and EvolveGCN), when they capture the dynamic patterns, they do not consider the dynamic interactions between item temporal features and user interest diffusion. Thus, they can't unify the two kinds of dynamic information as our model does. Besides, their GNN propagation methods are not suitable for our rising star scenario, which is the same as the above mentioned two static models.

5.3 Component-Wise Evaluation of RiseNet

To validate the contribution of each component of our proposed RiseNet, we conduct ablation studies on our key components. The ablation experiments are listed below:

- **Effect of User Interest in Diffusion.** RiseNet-r: A simplified version of RiseNet that does not include the graph

TABLE 3
Ablation study of RiseNet on four datasets

Method	Cate-1			Cate-2			Cate-3			Cate-4		
	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score
RiseNet	0.4072	0.6476	**0.5000**	0.2058	**0.4488**	**0.2822**	**0.1911**	0.3093	**0.2362**	0.1640	**0.4860**	**0.2453**
RiseNet-r	0.2851	**0.6571**	0.3977	0.1503	0.2047	0.1733	0.0870	0.3750	0.1412	0.1111	0.4783	0.1803
RiseNet-np	0.3964	0.6381	0.4891	**0.2239**	0.3543	0.2744	0.1556	**0.4124**	0.2260	**0.1771**	0.3505	0.2353
RiseNet-nm	0.3892	0.6190	0.4779	0.1963	0.3307	0.2463	0.1339	0.3505	0.1937	0.0717	0.3608	0.1197
RiseNet-nc	**0.4939**	0.3905	0.4362	0.1838	0.2677	0.2179	0.1387	0.3402	0.1970	0.1523	0.2523	0.1915
RiseNet-nd	0.2712	0.6095	0.3754	0.1932	0.2677	0.2244	0.1143	0.3196	0.1685	0.0803	0.4330	0.1355

neural network model (predict rising stars without the Taocode diffusion information).

- **Effect of PickGate.** RiseNet-np: The Model is trained without the PickGate. We get the graph representation by directly apply summation function to all user embeddings.
- **Effect of multitask framework.** RiseNet-nm: The Model is trained without the supervision of diffusion graph scale prediction task. (Graph Neural Network parameters are only updated by the loss of rising star prediction.)
- **Effect of the coupled mechanism.** RiseNet-nc: A simplified version of RiseNet that does not include the coupled mechanism: Fuse Gate (We use a simple concatenate function to replace the Fuse Gate), and the Diffusion Gate (We remove the influence probability in the graph propagation process).
- **Impact of capturing temporal evolving pattern.** RiseNet-nd: We evaluate the effect of our dynamic model structure. We simply concatenate the learnt graph representation at every time step, and then concatenate it with item features. After getting the final representation, we use it to predict the rising stars through a MLP.

We report the evaluation results in Table 3. It can be observed that RiseNet achieves the best performance compared with all ablated model versions in F1-score. The ablation experiment results can be interpreted to suggest the following:

1) From the comparison results with RiseNet-r, our model achieves superior performance (improves 48.0% in F1-score on average). It shows that the Taocode diffusion data that we introduced in the rising star prediction task has significantly helped us to identify rising stars with its contained user interest information.

2) The efficacy of PickGate. Without the PickGate in the graph neural network module, the performance of the model has weakened. It illustrates that by distinguishing the roles of different users, we retain more important information in the final graph representation that helps with rising star identification.

3) The effectiveness of the multitask framework. Without the supervision of graph diffusion scale prediction task on Graph Neural Network part of model, the F1-score is obviously lower than that obtained by the full version. It shows our multi-task mechanism significantly helps our model to leverage the user interest from the dynamic diffusion graphs by adding self-supervision.

4) The efficacy of the coupled mechanism. Removal of the coupled mechanism (FuseGate and DiffGate) yields poorer result compared with that achieved by the full ver-

sion. It illustrates the effectiveness of our coupled mechanism, which captures the dynamic interplay between temporal item information and user interest information in diffusion.

5) The effectiveness of the dynamic model structure. Without the dynamic model structure, the performance of the model worsens significantly, which indicates that capturing the temporal dependency of both user interest and item features can help us to recognize rising stars.

Based on the above ablation study results, we can now easily answer Q2 and Q3: user interest in Taocode diffusion greatly help us with rising star prediction and each component of our model RiseNet significantly contributes to our model on the target task.
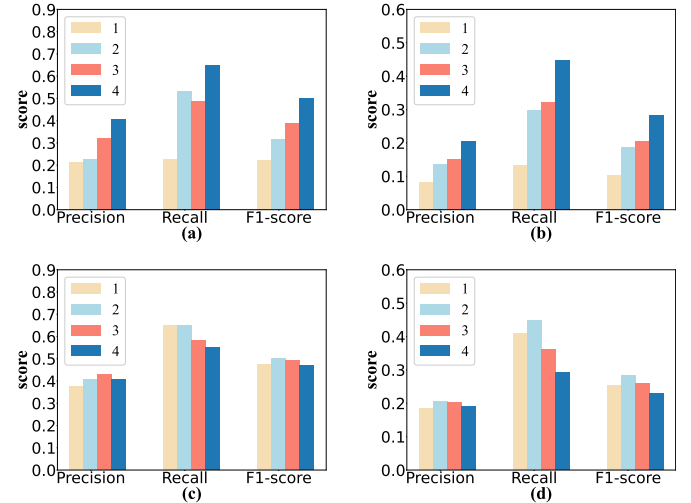

Fig. 5. Parameter analysis experiments of RiseNet. (a) The effect of time step length on cate-1 dataset. (b) The effect of time step length on cate-2 dataset. (c) The effect of GNN layer size on cate-1 dataset. (d) The effect of GNN layer size on cate-2 dataset.

5.4 Effect of dynamic information

To investigate the importance of the dynamic information, we examine the effect of utilized time steps' length on RiseNet. Specifically, length one means only the last timestep's information is utilized in rising star prediction (no dynamic information). We conduct experiments on the cate-1 and cate-2 datasets with various lengths (one to six weeks). From the experimental results in Figure 5, we can observe that with the increase of length of time steps, the model obviously performs better in F1-score. This illustrates that utilizing the dynamic information contained in items and Taocode data could help us to recognize rising stars.

5.5 Effect of GNN layer size

To answer Q5, we observe the effect of GNN layer size by conducting experiments with different GNN layer sizes: one to four. From the results in Figure 5(a) and (b), we can find that increasing layer size from one to two has little improvement on the model performance in F1-score. According to our observation on the Taocode users, more than 80% of users only have one-hop neighbor in the Taocode diffusion graph. Thus, increasing the propagation layer size cannot help us a lot to learn the user interest. Actually, applying layer size three and four even worsen the model performance compared to the result of applying layer size two, which is due to the over-smoothing problem.

6 RELATED WORK

In the interests of clarity, we first review sales prediction methods, then introduce works related to dynamic Graph Neural Networks.

Sales prediction. The rising star prediction problem can be regarded as a special sales prediction problem: predicting the potential sudden increase in the future sales. The sales prediction task has played an essential role in retailing. Traditionally, statistical methods are commonly used to predict sales; such methods include linear regression, moving average (MA), exponential smoothing [29], Bayesian analysis [30], and so forth. However, statistical methods fail to capture mutual non-linear dependencies among multidimensional time series data [31]. Thus, deep learning models like RNNs and CNNs have shown to be effective in sales prediction.

Borovykh etc. [19] present a deep convolutional WaveNet architecture as a solution to time series predicting (sales prediction is usually regarded as a time series prediction problem [32]). When multi-step predicting is required, the encoder-decoder architecture has been found to be helpful in real-life tasks [33] [34] [35]. LSTNet [20] utilizes CNN to capture short-term patterns and RNN to explore long-term dependence, which is frequently used in sales prediction. TCN [36] combines dilated convolutions with sequence modeling. Given millions of correlated time-series, DeepGLO [37] exploits both the local and global properties of time series and covariates. MTGNN [22] proposes a graph-based deep learning approach to exploit the inherent dependencies among multiple time series. However, most of the existing works focus on the long-term temporal prediction that has the periodicity or a regular developing trend. In our rising star prediction task, as the sales of rising stars change drastically in a short period of time, few existing works can keep their feasibility in capturing the rising signals. To address this problem, our work propose to utilize the user diffusion information to help the rising star prediction.

Beyond the sale prediction domain, there are some works that have addressed the user interest information [9][10]. [38] proposes a probabilistic model for predicting user purchase behavior with their tweets. [39] uses the user microblogging Information in cold-Start product recommendation problem. [40] utilizes tweets in product development. However, most of these works analyse user interest information by exacting several features from their social media content and ignore the dynamic information in user social network. To the best of our knowledge, we are the first to employ the user sharing graph into the sales prediction problem.

Dynamic Graph neural network. In our framework, we use the Graph Neural Networks to model the dynamic user diffusion graph. The GNNs have attracted significant research attention over the past decades [41] [23] [42] [43]. And as the development of static graph model, many researchers start to study the dynamic graphs [44][45][46]. Dynamic graph neural networks are designed to model the dynamic information as the graph evolving. Many real-world graphs are inherently dynamic, and their dynamic information helps to enhance the performance in many fields such as social network analysis and recommender systems. GCRN [25] is an extension of classical RNN on graph datasets, which exploits graphs with dynamic attributes to forecast moving MNIST data. DGNN [47] coherently captures the sequential information of edges, intervals between edges and information propagation to update node information. T-GCN [48] predicts traffic by combining GCN (where complex spatial dependence is learnt) with GRU (where temporal dependence is captured). TGAT [49] uses the self-attention mechanism and a time encoding technique to inductively infer node embeddings. The EvolveGCN [26] captures the dynamism of the graph sequence through using an RNN to evolve the GCN parameters. However, in our rising star scenario, the challenge lies on how to unify the dynamic information from both item and user diffusion graph information. This challenge makes the existing dynamic graph neural network unsuitable in this task. In order to solve this challenge, we propose to utilize the interplay between item and user interest with inspiration from observations. Different from the previous work, we add a special designed attention in the graph diffusion process with the utilization of dynamic information. Also, in each time step, we unify the item features and user diffusion information with the mutual dependency of them rather than simply concatenate them. This coupled mechanism enable our model to identify the rising stars more accurately.

7 CONCLUSION

In this paper, we propose to study a novel and practical problem: rising star prediction in online markets, aiming to find the promising items among the current low-turnover items. Rising star prediction benefits rational allocation of marketing resources, alleviation on unfair recommendation and balance of supply and demand. However, existing sales prediction methods are challenged by the contingency of the rising stars' sales trend. To address this, we introduce the user interest in diffusion to help with the target task and conduct a case study on Taocode diffusion data offered by Taobao. On the basis of the exploratory analysis on Taocode diffusion data, we validate the strong correlation between rising star identification and user interest diffusion. To this end, in order to leverage user interest and items features simultaneously to predict rising stars, a novel framework, named RiseNet, is proposed. RiseNet contains a coupled mechanism to model the dynamic interplay between user

interest and items, a special designed GNN based framework to quantify the user interest and a hierarchical structure to model temporal dependency. We prove the effectiveness of our model by obtaining better performance on Taobao real-world dataset compared with other state-of-art methods.

ACKNOWLEDGMENTS.

REFERENCES

[1] Y. Liu, K. Ge, X. Zhang, and L. Lin, "Real-time attention based look-alike model for recommender system," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019, pp. 2765–2773.

[2] X. Niu, B. Li, C. Li, R. Xiao, H. Sun, H. Deng, and Z. Chen, "A dual heterogeneous graph attention network to improve long-tail performance for shop search in e-commerce," in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2020, pp. 3405–3415.

[3] C. Anderson, *The long tail: Why the future of business is selling less of more*. Hachette Books, 2006.

[4] V. Ekambaram, K. Manglik, S. Mukherjee, S. S. K. Sajja, S. Dwivedi, and V. Raykar, "Attention based multimodal new product sales time-series forecasting," in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2020, pp. 3110–3118.

[5] Y. Qi, C. Li, H. Deng, M. Cai, Y. Qi, and Y. Deng, "A deep neural framework for sales forecasting in e-commerce," in *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 2019, pp. 299–308.

[6] J. Cummins and R. Mullin, *Sales promotion: How to create, implement and integrate campaigns that really work*. Kogan Page Publishers, 2010.

[7] J. Weng and B.-S. Lee, "Event detection in twitter," in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 5, no. 1, 2011.

[8] S. Gao, H. Pang, P. Gallinari, J. Guo, and N. Kato, "A novel embedding method for information diffusion prediction in social network big data," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 4, pp. 2097–2105, 2017.

[9] B. Hollerit, M. Kröll, and M. Strohmaier, "Towards linking buyers and sellers: detecting commercial intent on twitter," in *Proceedings of the 22nd international conference on world wide web*, 2013, pp. 629–632.

[10] X. W. Zhao, Y. Guo, Y. He, H. Jiang, Y. Wu, and X. Li, "We know what you want to buy: a demographic-based system for product recommendation on microblogs," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2014, pp. 1935–1944.

[11] R. B. Cialdini and R. B. Cialdini, *Influence: The psychology of persuasion*. Collins New York, 2007, vol. 55.

[12] G. R. Foxall, *Consumer behaviour analysis*. Taylor & Francis, 2002, vol. 3.

[13] B. Jin, C. Gao, X. He, D. Jin, and Y. Li, "Multi-behavior recommendation with graph convolutional networks," in *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval, SIGIR 2020, Virtual Event, China, July 25-30, 2020*, 2020, pp. 659–668.

[14] S. Brody, U. Alon, and E. Yahav, "How attentive are graph attention networks," *arXiv: Learning*, 2021.

[15] H. Gao and S. Ji, "Graph u-nets," in *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, ser. Proceedings of Machine Learning Research, K. Chaudhuri and R. Salakhutdinov, Eds., vol. 97. PMLR, 2019, pp. 2083–2092.

[16] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.

[17] G. E. Box and D. A. Pierce, "Distribution of residual autocorrelations in autoregressive-integrated moving average time series models," *Journal of the American statistical Association*, vol. 65, no. 332, pp. 1509–1526, 1970.

[18] G. P. Zhang, "Time series forecasting using a hybrid arima and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.

[19] A. Borovykh, S. Bohte, and C. W. Oosterlee, "Conditional time series forecasting with convolutional neural networks," *arXiv preprint arXiv:1703.04691*, 2017.

[20] G. Lai, W.-C. Chang, Y. Yang, and H. Liu, "Modeling long-and short-term temporal patterns with deep neural networks," in *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, 2018, pp. 95–104.

[21] Y. Qin, D. Song, H. Chen, W. Cheng, G. Jiang, and G. Cottrell, "A dual-stage attention-based recurrent neural network for time series prediction," *arXiv preprint arXiv:1704.02971*, 2017.

[22] Z. Wu, S. Pan, G. Long, J. Jiang, X. Chang, and C. Zhang, "Connecting the dots: Multivariate time series forecasting with graph neural networks," in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2020, pp. 753–763.

[23] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *ICLR*, 2017.

[24] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, "How powerful are graph neural networks?" *ICLR*, 2019.

[25] Y. Seo, M. Defferrard, P. Vandergheynst, and X. Bresson, "Structured sequence modeling with graph convolutional recurrent networks," in *International Conference on Neural Information Processing*. Springer, 2018, pp. 362–373.

[26] A. Pareja, G. Domeniconi, J. Chen, T. Ma, T. Suzumura, H. Kanezashi, T. Kaler, T. Schardl, and C. Leiserson, "Evolvegcn: Evolving graph convolutional networks for dynamic graphs," in *Proceedings of the AAAI Con-*

ference on Artificial Intelligence, vol. 34, no. 04, 2020, pp. 5363–5370.

[27] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga *et al.*, "Pytorch: An imperative style, high-performance deep learning library," in *NeurIPS*, 2019, pp. 8026–8037.

[28] M. Fey and J. E. Lenssen, "Fast graph representation learning with pytorch geometric," *arXiv preprint arXiv:1903.02428*, 2019.

[29] R. Hyndman, A. B. Koehler, J. K. Ord, and R. D. Snyder, *Forecasting with exponential smoothing: the state space approach*. Springer Science & Business Media, 2008.

[30] P. M. Yelland, S. Kim, and R. Stratulate, "A bayesian model for sales forecasting at sun microsystems," *Interfaces*, vol. 40, no. 2, pp. 118–129, 2010.

[31] F. M. Thiesing and O. Vornberger, "Sales forecasting using neural networks," in *Proceedings of International Conference on Neural Networks (ICNN'97)*, vol. 4. IEEE, 1997, pp. 2125–2128.

[32] T. Chen, H. Yin, H. Chen, L. Wu, H. Wang, X. Zhou, and X. Li, "Tada: trend alignment with dual-attention multi-task recurrent neural networks for sales prediction," in *2018 IEEE International Conference on Data Mining (ICDM)*. IEEE, 2018, pp. 49–58.

[33] W. Bao, J. Yue, and Y. Rao, "A deep learning framework for financial time series using stacked autoencoders and long-short term memory," *PloS one*, vol. 12, no. 7, p. e0180944, 2017.

[34] A. Sordoni, Y. Bengio, H. Vahabi, C. Lioma, J. Grue Simonsen, and J.-Y. Nie, "A hierarchical recurrent encoder-decoder for generative context-aware query suggestion," in *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, 2015, pp. 553–562.

[35] S. Bengio, O. Vinyals, N. Jaitly, and N. Shazeer, "Scheduled sampling for sequence prediction with recurrent neural networks," *arXiv preprint arXiv:1506.03099*, 2015.

[36] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," *arXiv preprint arXiv:1803.01271*, 2018.

[37] R. Sen, H.-F. Yu, and I. Dhillon, "Think globally, act locally: A deep neural network approach to high-dimensional time series forecasting," *arXiv preprint arXiv:1905.03806*, 2019.

[38] Y. Zhang, M. Shirakawa, and T. Hara, "A probabilistic model for predicting e-commerce purchases from emerging local and social interests," in *2020 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*. IEEE, 2020, pp. 345–350.

[39] W. X. Zhao, S. Li, Y. He, E. Y. Chang, J.-R. Wen, and X. Li, "Connecting social media to e-commerce: Cold-start product recommendation using microblogging information," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 5, pp. 1147–1159, 2015.

[40] A. A. Olad and O. F. Valilai, "Using of social media data analytics for applying digital twins in product development," in *2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*. IEEE, 2020, pp. 319–323.

[41] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Transactions on Neural Networks*, 2008.

[42] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," in *ICLR*, 2017.

[43] Z. Ying, J. You, C. Morris, X. Ren, W. Hamilton, and J. Leskovec, "Hierarchical graph representation learning with differentiable pooling," in *NeurIPS*, 2018.

[44] L. Zhu, D. Guo, J. Yin, G. Ver Steeg, and A. Galstyan, "Scalable temporal latent space inference for link prediction in dynamic social networks," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 10, pp. 2765–2777, 2016.

[45] L. Zhou, Y. Yang, X. Ren, F. Wu, and Y. Zhuang, "Dynamic network embedding by modeling triadic closure process," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, 2018.

[46] Y. Zuo, G. Liu, H. Lin, X. Guo, J.and Hu, and J. Wu, "Embedding temporal network via neighborhood formation," pp. 2857–2866, 2018.

[47] Y. Ma, Z. Guo, Z. Ren, J. Tang, and D. Yin, "Streaming graph neural networks," in *SIGIR '20: The 43rd International ACM SIGIR conference on research and development in Information Retrieval*, 2020.

[48] L. Zhao, Y. Song, C. Zhang, Y. Liu, P. Wang, T. Lin, M. Deng, and H. Li, "T-gcn: A temporal graph convolutional network for traffic prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 9, pp. 3848–3858, 2019.

[49] D. Xu, C. Ruan, E. Korpeoglu, S. Kumar, and K. Achan, "Inductive representation learning on temporal graphs," 2020.

Xuan Yang is currently a master student of the College of Computer Science and Technology, Zhejiang University, China. She received her bachelor's degree from the College of Computer Science and Technology, Zhejiang University. Her main research interests include data mining, graph modeling and computational social science.
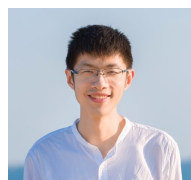
Yang Yang received his Ph.D. degree from Tsinghua University in 2016. He is an associate professor in the College of Computer Science and Technology, Zhejiang University. His main research interests include data mining and social network analysis. He has been visiting scholar at Cornell University and Leuven University. He has published over 40 research papers in major international journals and conferences including: KDD, WWW, AAAI, TKDE and TOIS.

Jintao Su is currently a master student of the College of Computer Science and Technology, Zhejiang University, China. He received his bachelor's degree from the College of Computer Science and Technology, Hangzhou Dianzi University. His main research interests include data mining and large scale graph modeling.

Yifei Sun is currently a master student of the School of Aeronautics and Astronautics at Zhejiang University, China. He received his BS degree in Computer Science and Technology, and BA degree in Japanese from Dalian University of Technology, in 2014. His current research interests include data mining and graph modeling.
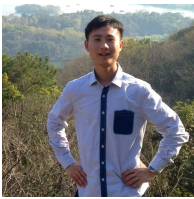
Shen Fan graduated from School of Computer Science and Engineering, University of Electronic Science and Technology of China(UESTC) in 2014. He is a staff engineer in Alibaba Group. His main research interests include social network mining and recommender systems. He is a member of CCF (China Computer Federation), and received the Elite Collegiate Award of CCF.

Zhongyao Wang received a B.Sc. in mathematics from the School of Mathematical Sciences, Peking University, Beijing, China in 2009. He joined the Data Department of Alibaba Group, Hangzhou, China in 2013, where he currently is a staff engineer. From 2010 to 2013, he worked in ASML (Brion) as a software engineer. His main research interests include data mining, graph learning, AutoML and recommendation. He worked on user profile and ID-mapping recent years, and currently leads a team responsible for an end-to-end automl system at Alibaba Group.
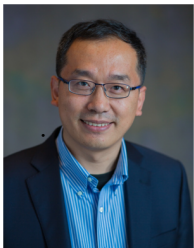
Jun Zhang holds a PhD from the University of Michigan (Ann Arbor) and bachelor from Tsinghua University at China. He currently leads the Data Assets and Algorithm team at Alibaba Data and Technology Department. Dr. Zhang is a known visionary leader in the industry applications of data science, cloud computing, and complex systems. He has published more than 40 scientific articles and research papers in the areas of complex network analysis, human computer interaction, and geo-spatial data science.
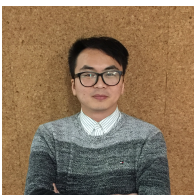
Jingmin Chen received his Ph.D. degree from Swiss Institute of Technology in Lausanne (EPF) in 2013. He is a senior staff engineer in Alibaba Group. His main research interests focus on spatial-temporal data mining, machine learning and transportation. He has published over 10 research papers in major internaltional journals and conferences.