Predicting Customer Value with Social Relationships via Motif-based Graph Attention Networks

Jinghua Piao, Guozhen Zhang, Fengli Xu, Zhilong Chen, Yong Li
Beijing National Research Center for Information Science and Technology (BNRist),
Department of Electronic Engineering, Tsinghua University, Beijing 100084, China.
Corresponding Email: liyong07@tsinghua.edu.cn

ABSTRACT

Customer value is essential for successful customer relationship management. Although growing evidence suggests that customers' purchase decisions can be influenced by social relationships, social influence is largely overlooked in previous research. In this work, we fill this gap with a novel framework — Motif-based Multi-view Graph Attention Networks with Gated Fusion (MAG), which jointly considers customer demographics, past behaviors, and social network structures. Specifically, (1) to make the best use of higher-order information in complex social networks, we design a motif-based multi-view graph attention module, which explicitly captures different higher-order structures, along with the attention mechanism auto-assigning high weights to informative ones. (2) To model the complex effects of customer attributes and social influence, we propose a gated fusion module with two gates: one depicts the susceptibility to social influence and the other depicts the dependency of the two factors. Extensive experiments on two large-scale datasets show superior performance of our model over the state-of-the-art baselines. Further, we discover that the increase of motifs does not guarantee better performances and identify how motifs play different roles. These findings shed light on how to understand socio-economic relationships among customers and find high-value customers.

CCS CONCEPTS

• **Information systems → Computational advertising**; **Social advertising**; • **Applied computing → Online shopping**; • **Computing methodologies → Modeling methodologies**.

KEYWORDS

Customer value prediction, graph neural networks, multi-view graphs.

ACM Reference Format:
Jinghua Piao, Guozhen Zhang, Fengli Xu, Zhilong Chen, Yong Li. 2021. Predicting Customer Value with Social Relationships via Motif-based Graph Attention Networks. In *Proceedings of the Web Conference 2021 (WWW '21), April 19–23, 2021, Ljubljana, Slovenia.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3442381.3449849

WWW '21, April 19–23, 2021, Ljubljana, Slovenia

ACM ISBN 978-1-4503-8312-7/21/04.
https://doi.org/10.1145/3442381.3449849

1 INTRODUCTION

As economics becomes predominantly service-based, companies are increasingly concentrating on managing good relationships with their customers and deriving sustainable revenue from these relationships [15]. Defined as forecasting the profits that each customer brings to the company in a future period, predicting customer value is essential and fundamental to customer relationship management (CRM) [11]. The ability to accurately predict customer value enables companies to build relationships with their customers based on their future worth rather than their past behaviors, and has been proven to be the key to developing effective personalized marketing campaigns [58].

Existing solutions adopt two mainstream approaches to predicting customer value, through qualitative models or quantitative models. The qualitative models usually base on marketing theories, and primarily focus on the behaviors of customers themselves. For example, the widely-accepted model RFM segments customer value based on recency, frequency, and monetary values [11]. As for quantitative models, Kumar et al. [33] utilize Markov processes to predict customer value, which demonstrates the importance of modeling the fine-grained customer historical behavior data. Further, researchers propose more complicated models, including traditional machine learning models [17, 56] and deep learning models [19]. All these models only consider individuals' behaviors, and do not take social relationships into account. However, recent research has demonstrated the significance of social relationships on customers' purchase decisions, and in turn, on customer value [61]. Thus, the lack of consideration on the effects of social influence arouses concerns on prior customer value prediction models.

To bridge this gap, in this paper, we aim to incorporate social network information to customer value prediction. To achieve this, key challenges as follows can emerge:

- **Unclear predictive signal of network structures.** Growing evidence suggests the influence propagated in social networks orients people's decision-making processes [2, 47, 61]. However, it is unclear whether and which social network structures contain useful signals for the customer value prediction task. As such, it is challenging to extract useful signals from social network structures.
- **Limited knowledge of higher-order information.** Prior work has demonstrated that social networks involve not only lower-order features at the node level, but also substantial non-trivial higher-order features at the level of small subgraphs [7, 41]. However, it is challenging to capture and utilize higher-order information in complex social networks,

and it is even harder to distinguish informative higher-order structures related to customer value.

- **Complex effects of customer attributes and social influence.** The effects of customer attributes and social influence on customer value can be dependent and entangled. For example, customer attributes, such as personality and occupations, can determine one's power to influence friends' purchase behaviors and susceptibility to friends' influence. Therefore, it is challenging to model such complex effects of customer attributes and social influence.

To address the above challenges, we propose a novel framework — Motif-based Multi-view Graph Attention Networks with Gated Fusion (MAG), where we jointly model the effects of customer attributes and friends' influence on customer value. Specifically, (1) to capture and utilize network information, especially higher-order information in complex social networks, we design a motif-based multi-view graph attention module. We leverage *motif*, the building block of complex networks, to construct motif-based multi-view graphs, and adopt attention mechanisms to fuse multi-view graph embeddings. (2) To model the complex effects of customer attributes and social influence, we design a gated fusion module, where we adopt two gate structures: one to model customers' susceptibility to social influence, and the other to model the dependency of customer attributes and social influence on customer value.

We evaluate our proposed framework on a large-scale real-world dataset collected from a leading social e-commerce platform in China, Beidian, along with an open dataset collected from a product review site, Epinions. On both platforms, customer value prediction is regarded as vital. Extensive experiments show that our model superiorly surpasses the state-of-the-art baselines on both datasets.

Furthermore, to better unleash the roles of higher-order information, we investigate the significance of different motifs. We discover that (1) even though more motifs contain diverse higher-order information, the increase of motifs does not guarantee better performances, and (2) in customer value prediction, effective information contained in different motifs varies significantly. Especially, it is observed that the "fully-connected" and the "down-linked mutual dyad" motifs are essential. The "fully-connected" motif indicates a close relationship and shows the capability to detect high-value communities. Besides, we observe that different from other social edges, the two unidirectional edges in the "down-linked mutual dyad" motif are mostly attributed to marketing behaviors, which uncovers the interplay between social influence and marketing efforts.

The contributions of this work can be summarized as follows:

- To the best of our knowledge, we are the first to leverage social network structures to model the social influence in customer value prediction. We demonstrate the importance of modeling network structures, especially higher-order structures, in customer value prediction.
- To predict customer value using social relationship information, we propose a framework — MAG, which jointly models customer attributes and social influence. Extensive experiments have demonstrated our proposed model's superior performance compared with the state-of-the-art baselines.
- Through an in-depth analysis of motifs, we discover that the "fully-connected" motif possesses the ability to detect high-value customers and the "down-linked mutual dyad" motif reveals communities' formation under the interplay between marketing efforts and social influence. These findings provide new insights into understanding socio-economic relationships among customers and excavating high-value customers.

2 PROBLEM FORMULATION AND MOTIVATION

To better explain our task and solution, we first introduce some necessary notations, and formally formulate the customer value prediction problem. Thereafter, we discuss our motivation to incorporate social relationships into the long-standing problem.

2.1 Problem Formulation

To properly formalize the task of incorporating social networks into customer value prediction, we first give a definition of social network $\mathcal{G}$: given n customers with k customer relations, we define a social network as a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of customers with $|\mathcal{V}| = n$, and $\mathcal{E}$ is the set of customer relations with $|\mathcal{E}| = k$. We denote the adjacent matrix of $\mathcal{G}$ as $\boldsymbol{A} \in \mathbb{R}^{n \times n}$.

Therefore, our task can be formulated as follows: given customer attributes and social networks, predict her/his customer value. Here we use the total amount of money or efforts that the customer will spend on the platform over the next period to measure customer value, which can reflect the value of the relationship with the customer. Moreover, customer attributes are extracted from customers' demographics and transaction history. Specifically, customers' demographics represent customers' basic personal information, including gender, age, zodiac, register time, and et al. Transaction history contains customers' total purchase value, frequency and intervals, along with return behaviors, such as total returned purchase value, times, and et al. The features extracted from customers' demographics and transaction history make up customer attributes. In addition, social networks are constructed based on social relations between customers, including friendship, trust relations, follower-followee relations, and et al.

2.2 Motivation

Many researchers are devoted to analyzing the complex structures of social networks [2, 4]. They discover that the social network structure is one of the strongest predictors for user behaviors [2, 4], leadership success [63], and future development [10]. In the area of CRM, social network structures have also been proved effective in predicting user engagement [42] and churn [36]. In this work, before considering incorporating social relationships into the task to facilitate the accurate customer value prediction, one fundamental hypothesis is that customer value is correlated with the customer's social network structures.

Thus, to validate this hypothesis, we conduct empirical data analyses on a large-scale dataset covering 44,995 customers with their social networks. To quantify social network structures, we adopt widely-accepted attributes describing the network properties in view of nodes, e.g., degree, coreness [6], pagerank score [43],

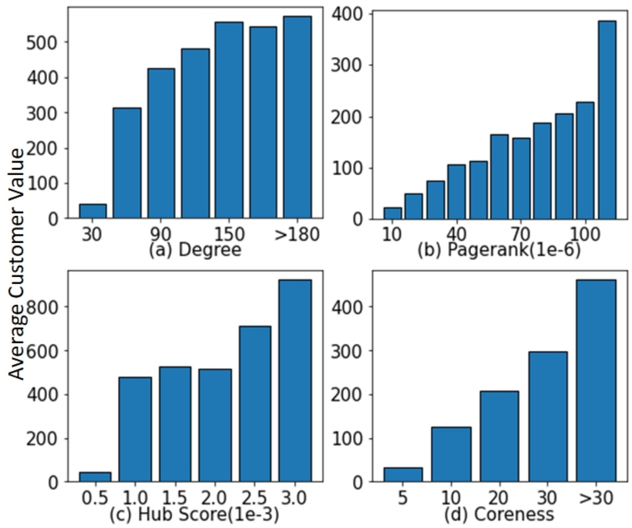

Figure 1: The correlations between customer value and network structures.

Model	Score	Feature Set	Importance
RF	38.02	Demographics	0.0856
		Recency	0.4920
		Frequency	0.0899
		Monetary Value	0.1967
		Return History	0.0165
RF-Network	33.79	Network Attributes	0.1193

Table 1: Comparison of performances between RF and RF-Network. Note that we use MAE to measure model performances.

hub and authority score (HITS) [31]. We analyze the relationships between these network attributes and customer value, where Figure 1 shows the correlations between customer value and (a) degree, (b) pagerank score, (c) hub score, and (d) coreness. From the figure, we can observe positive correlations between customer value and these network attributes. This indicates that customers with more friends (higher degree), higher popularity (higher pagerank score), higher hub scores, or higher corenesses, are more likely to have high customer value. These findings confirm our primary hypothesis that customer value is strongly correlated with social network structures and lay as the premise of the rest of this work.

To further examine whether these network attributes are beneficial for prediction, we incorporate them into the set of features in Random Forest model that is a traditional machine learning model widely-adopted in prior works on customer value prediction [17, 56]. Therefore, we obtain two models: one trained on customer attributes (RF) and the other trained on customer attributes and network attributes (RF-Network). As shown in Table 1, we discover that network attributes can improve the performances by 11.13% in terms of MAE. It indicates that social network structures are of great help to the prediction task. Meanwhile, the right two columns in Table 1 show features' importance scores generated by Random Forest model, which indicates that network attributes play a non-trivial role in prediction. To better utilize social network

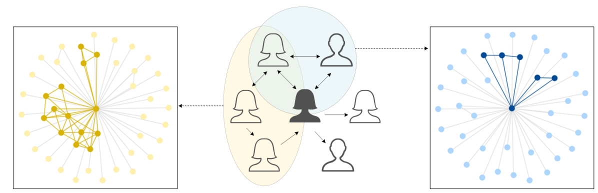

Figure 2: Illustration of higher-order structures captured by motifs in complex social networks.

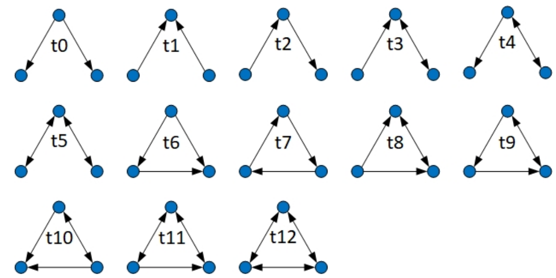

Figure 3: Illustration of thirteen three-node motifs.

information, we follow Qiu et al. [47], which finds that compared with these network attributes, graph neural networks (GNNs) show a superior ability to graph-structured modeling. This motivates us to consider leveraging GNNs, instead of mere network attributes, to exploit the potential of social networks. Besides, we further discuss the effectiveness of GNNs in Section 4.

Most researches on GNNs are devoted to modeling nodes' local connectivity [60]. However, considering the complex nature of social networks [28], it is insufficient to only consider lower-order structures, at the level of nodes, in complex social networks [7, 38]. Higher-order structures of complex networks, at the level of small subgraphs, provide a peek into the fundamental structures that control and mediate the behaviors of many complex systems [38, 46]. One approach to retaining higher-order structures in complex networks is introducing *motifs* into the framework. Motifs are special small subgraphs, and regarded as the building blocks of complex networks [38]. Motifs can capture high-order structures explicitly and effectively [34, 40, 49]. For example, as shown in Figure 2, we randomly sample one customer, i.e., the central node in the deep color, and her social networks. She has multiple friends, and these friends form small close groups. Motifs can easily capture higher-order structures, i.e., small groups in complex social networks.

To sum up, we validate our primary hypothesis: customer value correlates with social network structures. Moreover, we find out that motifs are able to capture higher-order structures and disentangle complex relationships in complex social networks. These observations motivate us to leverage social network structures, especially higher-order structures, to predict customer value.

3 METHOD

In order to incorporate social network structures, especially higher-order structures, into customer value prediction, we propose MAG, which jointly models the effects of customer attributes and their social network structures on customer value. Figure 4 shows the

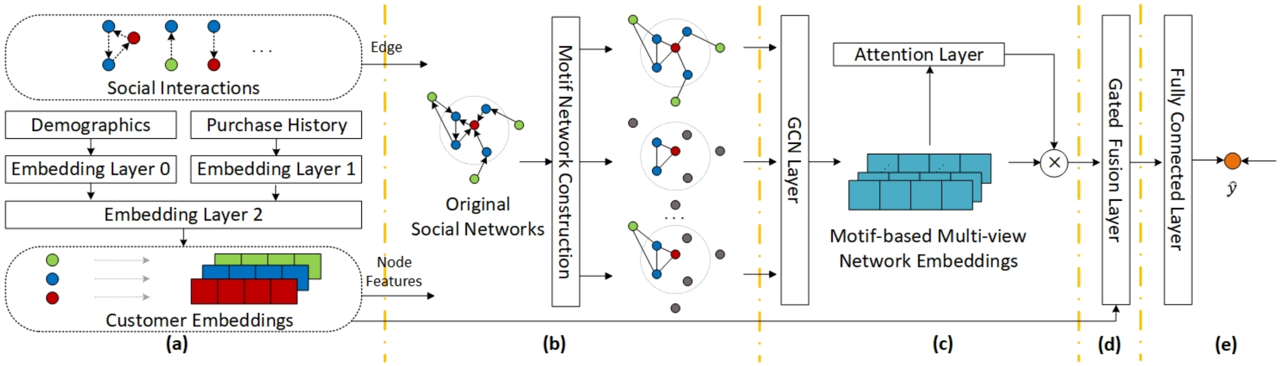

Figure 4: The architecture of MAG. (a) Inputs: Based on customers' demographics, transaction history and social interactions, we construct social networks and customer embeddings. (b) Motif network construction: Using motifs and social networks, we generate motif-based multi-view networks, motif networks for short. (c) Motif-based multi-view graph attention (MMA) module: After convolutional operation on motif networks, we apply attention mechanisms to fuse these multi-view graph embeddings. (d) Gated fusion layer (GF): To model the complex effects of customer attributes and social influence, we design two gate structures in the module. (e) Outputs: Multiple fully-connected layers are adopted to map the embeddings to the predicted value. Finally, we compare the predicted value with the ground truth.

architecture of MAG. Specifically, (a) Inputs construct social networks based on interaction records and encode customer attributes into embeddings. (b) Motif network construction module generates motif-based multi-view graphs that explicitly retain various higher-order structures in the original social networks. (c) Motif-based multi-view graph attention (MMA) module learns the corresponding embeddings from these graphs and adopts the attention mechanism to fuse these embeddings, which allows it to make the best use of various higher-order information in complex social networks. (d) Gated fusion layer (GF) utilizes two gate structures to model the complex effects of customer attributes and social networks. (e) Outputs map the fused embeddings generated by GF into the predicted value. The details of the above layers of MAG are introduced as follows.

3.1 Motif Network Construction

Inspired by the work of Lee et al. [34] , we adopt a motif-based matrix formulation to construct motif-based customer social networks (motif networks for short). Given a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and a set of N different motifs $\mathcal{M} = \{\boldsymbol{M_0}, \boldsymbol{M_1}, \ldots, \boldsymbol{M_{N-1}}\}$, we can construct N different motif-based adjacent matrices $\mathcal{A} = \{\boldsymbol{A_0}, \boldsymbol{A_1}, \ldots, \boldsymbol{A_{N-1}}\}$, where

$$(\boldsymbol{A_n})_{i,j} = \begin{cases} 1 & i = j, \\ 1 & i, j \text{ are in the same } \boldsymbol{M_n}, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

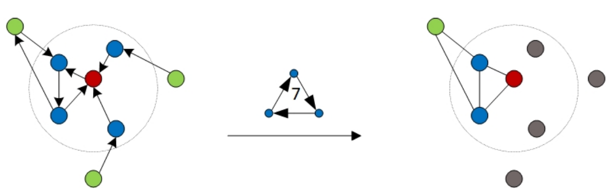

Figure 5: Illustration of motif network construction.

Figure 5 shows an example of Motif t7 Network construction. Specifically, the links in the original network (left) that can match Motif t7 are preserved in the new Motif t7 Network (right), and the others are discarded. Therefore, some adjacent nodes are disconnected in motif networks, which indicates that the neighbors are refined by the process [34]. What's more, it is through nodes' adjacency in the motif network that we imply specific subgraph structures, i.e., motifs, in the original network. Therefore, the motif network is undirected. To sum up, motif network construction helps capture higher-order interactions among nodes in the graph and refine the neighbors. The analysis on the effectiveness of motif networks and the function of motifs in complex social networks is left for Section 5.

3.2 Motif-based Multi-view Graph Attention

The construction of motif networks generates N different motif-based adjacent matrices $\mathcal{A} = \{\boldsymbol{A_1}, \boldsymbol{A_2}, \ldots, \boldsymbol{A_N}\}$. In order to obtain network embeddings of each motif network, we perform graph convolutional operations [30] on them respectively:

$$\boldsymbol{X_n^{l+1}} = \sigma(\boldsymbol{D_n^{-1/2} A_n D_n^{-1/2} X_n^l \Theta^l}), \tag{2}$$

where $\sigma(\cdot)$ is an activation function, such as ReLU; $\boldsymbol{X_n^l}$ represents the network embeddings in the l−th layer from motif network $\boldsymbol{A_n}$, $(\boldsymbol{D_n})_{ii} = \sum_{j=0}(\boldsymbol{A_n})_{ij}$ is its diagonal degree matrix and $\boldsymbol{\Theta^l}$ are learnable parameters.

Considering not all motif networks are equally important [34], we adopt attention mechanisms to fuse N multi-view graph embeddings, which can ensure that high weights are auto-assigned to informative ones:

$$X_{infl} = \sum_{n=0}^{N-1} \alpha_n \cdot \boldsymbol{X_n^{l+1}}, \tag{3}$$

where X_{infl} represents the fused motif-based multi-view graph embeddings and α_n is the attention coefficient of Motif-n Network. The attention coefficients can be calculated as follows:

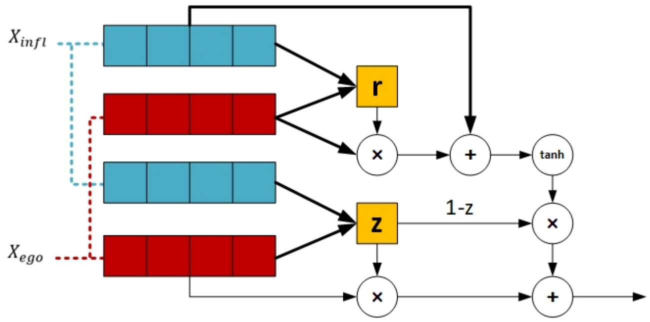

Figure 6: Gated fusion layer, where the thick lines represent that the input will be applied a linear transformation while the thin lines denote that the input is directly linked to the output.

$$\alpha_n = \frac{\exp(e_n)}{\sum_{i=0}^{N-1} \exp(e_i)}, \tag{4}$$

$$e_n = \boldsymbol{W}_c \tanh(\boldsymbol{W}_e \boldsymbol{X}_n^{l+1}), \tag{5}$$

where $\boldsymbol{W}_c$ and $\boldsymbol{W}_e$ are randomly-initialized learnable parameters. $\boldsymbol{W}_c$ serves as a learnable contextual vector in this case.

3.3 Gated Fusion Layer

Social influence also closely relates to receivers' characteristics, for example, their susceptibility to influence [3, 8, 39]. Because MMA focuses on capturing higher-order structures in social networking without explicitly considering the effects of customer attributes on the process of message-passing, we design GF to depict the complex effects of customer attributes and social influence on customer value. Inspired by the design of GRU [22], GF contains two gate structures: the $\boldsymbol{r}$ gate and the $\boldsymbol{z}$ gate. Figure 6 shows the structure of GF, where the inputs are the embeddings of customer attributes $\boldsymbol{X}_{ego}$ introduced by the residual path from the input layer and the network embeddings $\boldsymbol{X}_{infl}$ from MMA. The inputs $\boldsymbol{X}_{ego}$ and $\boldsymbol{X}_{infl}$ control both the $\boldsymbol{r}$ and the $\boldsymbol{z}$ gate. Specifically, the $\boldsymbol{r}$ gate is designed for carving out the customer's susceptibility to influence, which can be formulated as follows:

$$\boldsymbol{r} = \sigma(\boldsymbol{W}_{ego,r}\boldsymbol{X}_{ego} + b_{ego,r} + \boldsymbol{W}_{infl,r}\boldsymbol{X}_{infl} + b_{infl,r}). \tag{6}$$

Here, a higher $\boldsymbol{r}$ represents less susceptibility to influence, or in other words, more susceptibility to insist on herself/himself. Therefore, the social influence, denoted by $\hat{\boldsymbol{X}}_{infl}$, can be calculated as follows:

$$\hat{\boldsymbol{X}}_{infl} = \tanh(\boldsymbol{W}_{infl,n}\boldsymbol{X}_{infl} + b_{infl,n} + \boldsymbol{r} \odot (\boldsymbol{W}_{ego,n}\boldsymbol{X}_{ego} + b_{ego,n})). \tag{7}$$

Further, to jointly model the complex effects of customer attributes and social influence on customer value, we design the $\boldsymbol{z}$ gate:

$$\boldsymbol{z} = \sigma(\boldsymbol{W}_{ego,z}\boldsymbol{X}_{ego} + b_{ego,z} + \boldsymbol{W}_{infl,z}\boldsymbol{X}_{infl} + b_{infl,z}). \tag{8}$$

Through the $\boldsymbol{z}$ gate, we fuse the embeddings of customer attributes $\boldsymbol{X}_{ego}$ and social influence $\hat{\boldsymbol{X}}_{infl}$, which can be formulated as follows:

$$\boldsymbol{X}_{fused} = (1 - \boldsymbol{z}) \odot \hat{\boldsymbol{X}}_{infl} + \boldsymbol{z} \odot \boldsymbol{X}_{ego}, \tag{9}$$

where $\boldsymbol{X}_{fused}$ denotes the fused embeddings, $\boldsymbol{z}$ represents the weights of customer attributes and $1 - \boldsymbol{z}$ represents the weights of social influence. Note that in above equations, $\boldsymbol{W}$ and b are learnable parameters, $\sigma(\cdot)$ is an activation function, e.g., sigmoid and $\odot$ represents Hadamard product.

To summarize, GF utilizes two gate structures to model the complex effects of customer attributes and social influence. In addition, the residual path makes it easier to train customer embedding layers and it also mitigates noisy information from the increasing receptive field [35].

3.4 Inputs, Outputs and Loss Functions

Inputs: The input layer consists of two key components: social network construction and customer embeddings. To represent the social relationships among customers, we adopt a simplified binary indicator: observed social interactions between the two customers are represented by a directed edge from the sponsor to the follower with the weight of 1. To represent customers, we learn from past research which indicates that purchase recency, frequency and monetary value are three valuable indicators in the customer value prediction task [11, 17, 26, 56]. Thus, we extract the node feature set not only from customers' demographics, but also from transaction history. This allows our feature set to cover the major dimensions of customer attributes, especially purchase recency, frequency and monetary value. Further, to encode customer attributes $\boldsymbol{X}_{ego}$, we utilize the embedding layers shown in Figure 4(a).

Outputs: The prediction layer takes the fused embeddings $\boldsymbol{X}_{fused}$ from GF as inputs to predict customer values. Here, we use multiple fully-connected layers to map the fused embeddings $\boldsymbol{X}_{fused}$ into the predicted value $\hat{y}$:

$$\boldsymbol{h}^{l+1} = \sigma(\boldsymbol{W}_h^l \boldsymbol{h}^l + b_h^l), \tag{10}$$

where l represents the l-th fully-connected layer; $\boldsymbol{W}_h^l$ and b_h^l are learnable parameters; $\boldsymbol{h}^l$ is the hidden representation in the l-th layer, $\boldsymbol{h}^0 = \boldsymbol{X}_{fused}$. Note that the regressor can be substituted by others, for example, lightGBM regressor [29].

Loss Functions: To minimize the absolute difference between the ground truth and the predicted value, we adopt mean absolute error (MAE) as the loss function, which is formulated as:

$$\mathcal{L} = \frac{1}{N} \sum_{n=0}^{N-1} |\hat{y} - y| + \lambda \sum ||\Theta||_2, \tag{11}$$

where N is the batch size, $\hat{y}$ and y are the predicted value and the ground truth respectively. To avoid over-fitting, we apply l_2 regularization to the parameters.

4 EXPERIMENTS

To demonstrate the effectiveness of our proposed framework, we conduct extensive experiments on two real-world datasets. Specifically, in this section, we first describe the experimental setup, and then evaluate our proposed model in comparison with the state-of-the-art baselines. Finally, to further verify the performances, we conduct an ablation study of MMA and GF.

Dataset	#Customers	#Links	Avg. Deg.	Avg. Clus. Coef.
Beidian	44,995	146,812	4.1506	0.3027
Epinions	10,956	143,077	26.1184	0.1198

Table 2: Basic information of social networks.

4.1 Experimental Setup

4.1.1 Dataset. We evaluate our model based on two real-world datasets: Beidian and Epinions. The Beidian dataset is collected from one of China's leading social e-commerce platforms — Beidian[1]. Leveraging social relationships for item recommendations and promotions [13, 14, 20, 27, 61], the emerging form of social e-commerce offers an opportunity to render transaction data and social interactions without extra efforts to collect and link them together, where customer value is reflected in the total amount money that the customer spend on the platform over the next period. This dataset contains 44,995 customers, and their fine-grained customers' demographics, transaction history and social interactions in 2019. The large scale and good quality of it enable us to validate the idea of predicting customer value with social relationships. Following prior works [56], we divide the one-year-long dataset into four quarters chronologically. Specifically, we train with features up to the end of Quarter 1, and thereby predict customer value in Quarter 2. We adopt an out-of-time validation dataset that are shifted one quarter later, i.e., we tune the hyper-parameters through the prediction of customer value in Quarter 3. Finally, we take the features up to the end of Quarter 3 as the input to predict customer value in Quarter 4, which makes up the following evaluation results. Note that customer value is the only information extracted from the data of Q4, which is adopted as the predicted targets of Q3. Based on social interactions among 44,995 customers, their social networks are constructed, where the relationship is friendship. The basic information of the social networks is shown in Table 2, where the average degree reflects customers' average number of friends and the average clustering coefficient reflect the density of social ties. Note that the degree and clustering coefficient are averaged among all customers.

The other is an open dataset collected from a product review site Epinions [52, 53], where customer value is reflected by the number of product reviews a customer contributes to the platform. This dataset contains customers' historical rating records and trust relations between customers. Similarly, we divide the dataset from 2000 to 2003 into four parts by year chronologically. We train with features up to the end of 2000 and predict customer value in 2001. Then we tune the hyper-parameters by predicting customer value in 2002. Finally, models are evaluated on the data of features up to the end of 2002 and labels of customer value in 2003. Based on trust relations between customers, we construct the social networks, of which the basic information is shown in Table 2.

4.1.2 Baselines. In order to validate the performances of our model, we compare it with three lines of state-of-the-art baselines, including traditional machine learning (ML) models (Random Forest [17, 56], Support Vector Regression [23]), network embeddings (DeepWalk [44], LINE [54], WALKLETS [45]), and graph neural networks (GNNs) (GCN [30], GAT [57], Meta-GNN [51]). These methods are introduced as follows:

[1]https://www.beidian.com

- Random Forest [17, 56]: An ensemble learning method based on decision trees. Existing methods mostly take traditional features, such as demographics, purchase frequency, recency and monetary value as the input. Here, we conduct experiments based on three feature sets: (1) only traditional features (RF), (2) traditional and network features (RF-Network), and (3) traditional, network and motif features (RF-Motif).
- Support Vector Regression (SVR) [23]: A traditional supervised regression model. Note that we adopt the radial basis function kernel here. The input contains only traditional features.
- DeepWalk [18, 44]: An unsupervised network embedding method based on random walk and skip-gram algorithms. It is widely used to learn latent representation of nodes for downstream tasks.
- LINE [18, 54]: An unsupervised method for embedding very large information networks into low-dimensional vector space.
- WALKLETS [18, 45]: An upgrade version of DeepWalk that introduces skipping into the process of random walks. The method learns multi-scale representations of nodes in networks explicitly, which allows it to retain higher-order information to some extent.
- GCN [30]: A current widely-accepted variant of graph neural networks. Here, the node features are only traditional features and the input graph is the original social network.
- GAT [57]: A state-of-art GCN method by leveraging masked self-attentional layers to address the shortcomings of prior GCN methods. Note that the node features are only traditional features, and the input graph is the original social network.
- Meta-GNN [51]: A current graph neural network framework that explicitly utilizes higher-order network structures. Note that the node features are only traditional features, and the input graphs are all thirteen three-node motif networks.

4.1.3 Motif Selection. Motifs, i.e., small subgraphs in complex networks, are various. Among motifs, the ones with a few nodes, especially three-node motifs, have been proved to have the capability of controlling and mediating the behaviors of many complex systems [38], on which many theories are established [28, 38, 59]. Following prior literature in network science, we include all thirteen three-node motifs (see Figure3) as candidates. It is worth mentioning that the attention mechanism in MMA can auto-assign high weights to significant motifs, which enables MAG to have the ability to tolerate minor redundant or misleading motifs.

4.1.4 Metrics, Reproducibility and Ethical Considerations. To evaluate the performances, we adopt three widely-adopted metrics in regression tasks: mean absolute error (MAE), mean absolute percentage error (MAPE) and symmetric mean absolute percentage error (SMAPE) [37]. Specifically, we include MAE for its robustness to outliers, MAPE and SMAPE for their ability to interpret relative errors.

We implement our experiments with Pytorch, where hyper-parameters of GNNs are reported as follows: For the Beidian dataset, the output sizes of embedding layers in the input module are fixed to 32, 32 and 32; the number of GCN/GAT layers is 2, where the

Category	Models	Beidian			Epinions		
		MAE	MAPE	SMAPE	MAE	MAPE	SMAPE
Traditional ML	SVR	37.69	2.52	1.01	1.87	0.72	0.88
	RF	38.02	1.40	0.90	2.68	1.72	0.86
	RF-Network	33.79	1.31	0.90	2.64	1.67	0.86
	RF-Motif	32.76	1.30	0.90	2.62	1.65	0.86
Network Embedding	DeepWalk	24.34	0.17	0.16	1.37	0.25	0.28
	LINE	24.24	0.16	0.16	1.36	0.25	0.29
	WALKLETS	22.19	0.17	0.17	1.36	0.24	0.28
GNNs	GCN	24.25	0.16	0.16	1.54	0.27	0.36
	GAT	24.29	0.16	0.16	1.52	0.26	0.41
	Meta-GNN	24.45	0.17	0.18	1.43	0.24	0.30
Ours	MAG	**20.57****	**0.15****	**0.15****	**1.33****	**0.21****	**0.24***

Table 3: Overall performances in MAE, MAPE and SMAPE. We repeat 20 experiments and perform one-sided Welch's t-tests, where (**) indicates p<.01 and (*) p <.05.

output sizes are set as 32; the number of fully-connected layers are 4, where the output sizes are 32, 16, 8, 1, respectively. For the Epinions dataset, the number of embedding layers is 1 because the dataset only contains customers' historical behaviors, and the output size is set as 8. The number of GCN/GAT layers is 2, where the output sizes are 8. The number of fully-connected layers are 2, where the output sizes are 8 and 1. We adopt ADAM optimizer to train the models. Besides, we tune the learning rate, dropout rate and regularization parameters by grid searching. For network embedding methods, we pre-train the models and fix the outsize size to 32 (Beidian)/8 (Epinions). Then we concatenate customer embeddings with network embeddings and feed them into the 4 (Beidian)/2 (Epinions) fully-connected layers mentioned above. All the models trained on the server with two core CPU (AMD Ryzen Threadripper 2990WX) and four GPUs (NVIDIA GeForce RTX 2080Ti * 4). The implementation code of our model and detailed information are available at https://github.com/tsinghua-fib-lab/MAG-Customer-Value-Prediction.

Procedures are carefully taken to address the ethical issues. Data collection is under the Terms of Service for Beidian including consent for research studies. All potential personal identifiers are replaced with anonymous hashcodes to ensure data with users' privacy are sanitized. We store data and train models in an offline server where access is only limited to authorized researchers bound by confidentiality agreements.

4.2 Overall Performance

The overall performances of our baselines are reported in Table 3. From the table, we have the following observations and conclusions.

- Our model constantly outperforms all the baselines in terms of all the metrics. Notably, compared with the state-of-the-art baselines, our proposed model achieves relative performance gains of 7.30%, 6.25%, 6.25% on the Beidian dataset, and 2.2%, 12.5%, 14.3% on the Epinions dataset in MAE, MAPE and SMAPE, which corroborates the effectiveness of our proposed methods.
- Comparing the performances of RF, RF-Network and RF-Motif, we find out that the utilization of network structures, especially higher-order structures captured by motif networks, is beneficial.
- The comparison of RF-Network and GNN-based methods (including our model) demonstrates that network attributes, e.g., degree, coreness and pagerank score, reserve limited information in social networks, while GNN-based methods can make full use of neighbors' attributes to infer customer value.
- We observe that compared with other network embedding methods, WALKLETS [45] has a relatively good performance. The result can be attributed to the design of skipping in the process of random walk, which allows the model to retain higher-order information to some extent. This reconfirms the effectiveness of higher-order information.
- Our model's superior performances over network embedding methods and GNNs imply that higher-order information captured by our MMA module can greatly help customer value prediction.

4.3 Ablation study

Our proposed model consists of two key components, namely MMA and GF. To analyze the contributions of each component, we conduct an ablation study on variants of our model on the Beidian dataset. Here, variants include the model without either MMA or GF (GCN), the one with only MMA (Ours-MMA), the one with only GF (Ours-GF), and our complete model with both MMA and GF (MAG). The results are shown in Table 4. We observe that the introduction of MMA reduces MAE by 11.13% and GF by 4.24%, indicating that the two components contribute to the performance improvement. In addition, MAG with the two components reduces MAE, MAPE and SMAPE by 15.18%, 6.25% and 6.25%, respectively.

Model	MAE	MAPE	SMAPE
GCN	24.25	0.16	0.16
Ours-MMA	21.55	0.15	0.17
Ours-GF	23.22	0.16	0.16
MAG	**20.57**	**0.15**	**0.15**

Table 4: Ablation study on Beidian dataset.

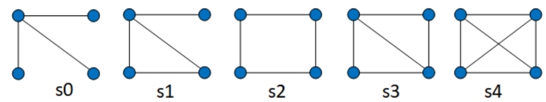

Figure 7: Five four-node motifs added into MAG.

Model	MAE (GAIN)	MAPE (GAIN)	SMAPE (GAIN)
Ours-orig	23.22 (-)	0.16 (-)	0.16 (-)
Ours-3motif	20.57 (**2.65**)	0.15 (**0.01**)	0.15 (**0.01**)
Ours-4motif	20.26 (0.31)	0.16 (-0.01)	0.14 (**0.01**)

Table 5: Performances of models with three-node-directed and four-node-undirected motifs on Beidian dataset.

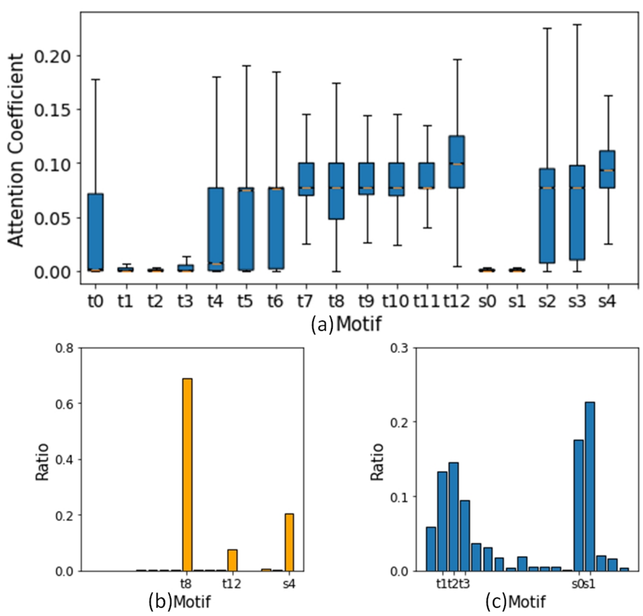

Figure 8: (a) Distribution of attention coefficients among motif networks. Ratio of nodes that assign (b) the highest, (c) the lowest attention weights to the corresponding motif network on Beidian dataset.

This indicates that the combination of the two components can effectively reduce both absolute and relative errors. To sum up, MAG performs the best among all the variants, and the two components can generally improve the performances.

5 IN-DEPTH ANALYSIS OF MOTIFS

To better understand the performance of MAG, we make an in-depth analysis of *motifs*, the essential higher-order information extractor. Past research has demonstrated that motifs are functional and informative in complex networks [38]. The design of MMA offers us a lens into the function of motifs in social networks. Therefore, we conduct extensive experiments on the Beidian dataset to answer: Do more motifs guarantee better performances? How do different motifs contribute to customer value prediction? Why are specific motifs informative or uninformative?

5.1 Motifs: Unequally Informative

In this work, we capture higher-order information by leveraging motifs to reconstruct the networks. Existing works reveal various motifs play different roles in complex networks, e.g., feed-forward loop in neurons, three-node feedback loop in electronic circuits [38]. However, most of them are limited to three-node motifs and few include multiple-node motifs because triadic structures have been well explained in sociology and the computational complexity of multiple-node motifs is high [25, 50]. They do not give a credible reason for whether we need multiple-node motifs and why. To fill this gap, we conduct the following experiments to answer whether more motifs can guarantee a better performance and why.

Intuitively, more motifs contain more diverse higher-order information, whereas it is unknown whether more diverse structural information can guarantee a satisfactory performance gain. To examine whether more motifs lead to better performances, we analyze the model performances with the addition of five four-node motifs (see Figure 7). Here, Ours-orig denotes the model with only the original social networks. Ours-3motif adds three-node motif networks to Ours-orig, and Ours-4motif further adds four-node motif networks to Our-3motif. The results are shown in Table 5.

From Table 5, we discover that incorporating three-node-directed motifs improves the performances greatly (11.41% in MAE), while further adding four-node-undirected motifs only leads to slight improvements (1.51% in MAE). This indicates that appropriate incorporation of certain motifs can contribute to better performances, but the performance improvements brought by extra motifs can diminish when the scale of motifs increases. One possible explanation is that more motifs introduce extra parameters, which is more likely to induce overfitting.

To better explain the phenomenon, we further analyze the attention coefficients in the motif-based multi-view graph attention layer. To focus only on motifs, we remove GF in the following experiments to avoid disturbance and compute the attention coefficients among different motif networks (see Figure 8(a)). Figure 8(b)(c) show the ratio of nodes that assign the highest/lowest attention weights to each motif network to all nodes, which measures the utility of the corresponding motif for the overall prediction task. We observe that varying attention weights are assigned to different motifs: Motif t8, t12 and s4 stand out while Motif t1, t2, t3, s0 and s1 receive relatively low attention weights. One possible reason for different weights is that in contrast to the performance improvements brought by the informative motifs, the introduction of the motifs with low weights might make limited contributions to or even harm the performances. To confirm the intuition, we conduct an experiment by removing Motif t1, t2, t3, s0 and s1. It is observed that the removal of the less informative motif networks does not hurt the performances in terms of MAE (0.7% improved), MAPE and SMAPE (unchanged).

In summary, more motifs cannot guarantee better performances, and motifs are NOT equally informative. Informative motifs can effectively improve the performances, while the introduction of uninformative ones might impair the model and waste extra computational resources.

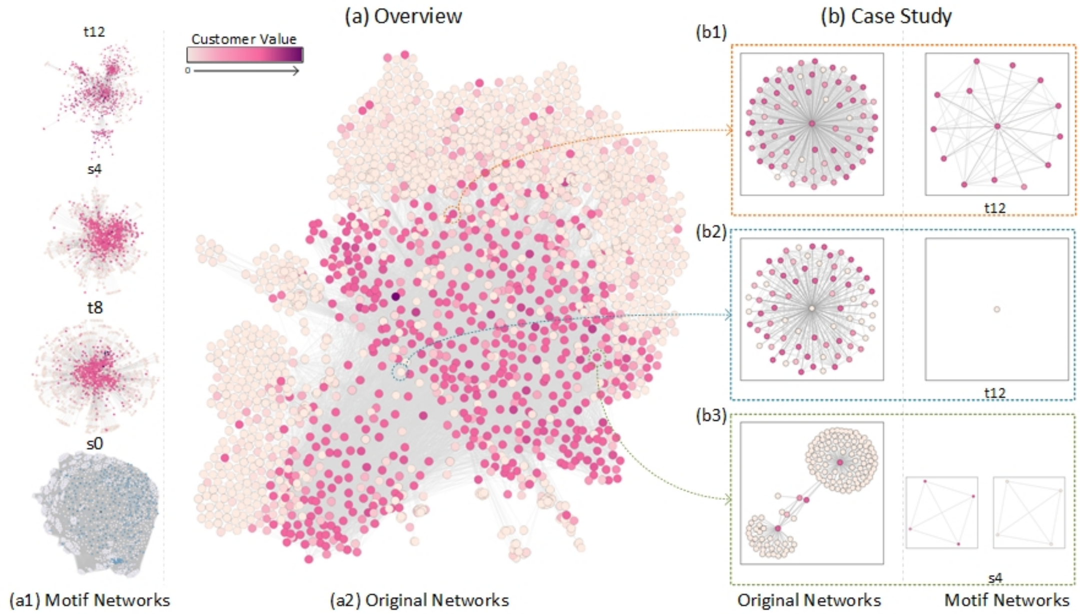

Figure 9: Comparison between the original network and the motif networks. We sample a small subgraph for illustration. (a) Overview: In the original network, customers with different values are mixed; while in Motif t12, s4 and t8 Networks, it is easier to distinguish customers. In addition, Motif s0 Networks introduce exponentially-large numbers of redundant links. (b) Case study: In the orange and the blue boxes, the two nodes have different customer values but similar distributions of neighbors in the original network. However, in the motif networks, (b1) the node has similar high value neighbors and (b2) the one is isolated. In the green box (b3), Motif s4 separates customers into high-value and low-value corps.

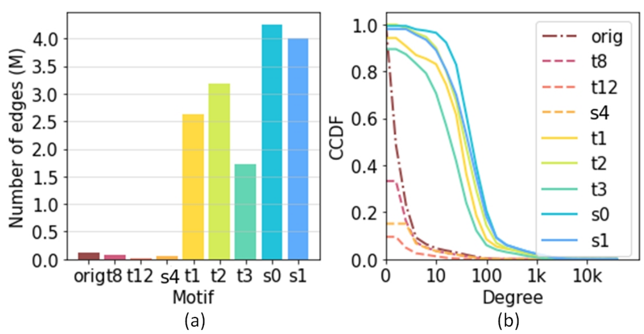

Figure 10: Statistics of motif networks: (a) the number of edges, (b) the distribution of degrees.

5.2 Explanation: Informative or Not

In this section, we aim to explain why Motif t8, t12 and s4 are informative, while Motif t1, t2, t3, s0 and s1 are less informative. As shown in Figure 3, the degrees of nodes in Motif t8, t12 and s4 are all greater than 1, which represents a relatively strong connection among the three nodes (or the four nodes). However, the nodes in Motif t1, t2, t3, s0 and s1 are sparsely connected. The statistics of the corresponding motif networks are shown in Figure 10. We discover that these informative motifs can selectively filter out multiple edges and efficiently select node neighbors. However, the uninformative ones introduce an exponentially-large number of extra edges and neighbors. These results indicate a critical function of informative motifs is extracting useful connections.

To further explore how the extraction of useful connections is functioning, we make an in-depth analysis of Motif t8, t12 and s4. As shown in Figure 9, customers with different values are mingled in the original network. However, in the informative motif networks, especially Motif t12/s4, communities with high values are extracted. In this way, informative motifs serve the function of refining connections between nodes.

Case studies further supplement the investigation on what role these informative motifs play in complex social networks. In Figure 9 (b1) and (b2), the two central nodes are customers who share similar neighbors in the original social networks but different customer values. However, in Motif t12 Network, the node in the orange box (b1) is surrounded by similar high-value (deep color) neighbors, while the one in the orange box (b2) is isolated. Similar cases are also spotted for Motif s4. As illustrated in Figure 9 (b3), Motif s4 separates complex structures into high-value corps and low-value corps. As such, Motif t12 and s4 refine the neighbors and partition complex networks into small homogeneous graphs, which can effectively facilitate the learning of network embeddings. This is in line with previous works, where the special "fully-connected" structures of Motif t12 and s4 have also been highlighted in various complex networks [7, 32, 38].

However, the function of Motif t8 is still under-explored. As shown in Figure 11, Motif t8 is a symmetric graph with two important positions, denoted by *Position a* and *Position b*. We observe that 84.3% of customers in *Position a* have succeeded in recommending an item to those in *Position b*. This indicates that the link between *Position a* and *Position b* is distinguished from other links that are

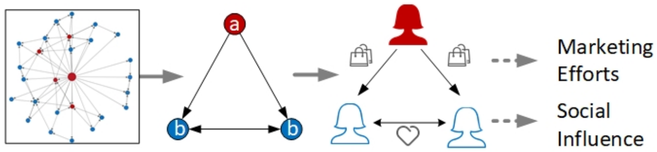

Figure 11: Illustration of the function of Motif t8.

mainly driven by social influence. Instead, the mechanism of marketing efforts can be a plausible reason for the recommendation behaviors from *Position a* to *Position b* [16, 55]. Furthermore, Motif t8 reveals the formation of buying groups in social e-commerce: *Position a* provides recommendation service to *Position b*, and if *Position b* finds it satisfactory, she/he invites friends to join them. Thus, the vital role of Motif t8 in the socio-economic networks accounts for its benefits to the prediction.

In summary, we discover informative motifs, especially Motif t12 and Motif s4, possess the great ability to detect high-value communities in complex social networks. Besides, Motif t8 reflects the unique socio-economic characteristics embedded in the network. These findings offer a new peek into understanding the interplay between social influence and economic factors.

6 RELATED WORK

In this section, we position our work in the related literature from the following aspects: customer value prediction (the long-standing marketing problem our work focuses on), graph neural networks (the technical background our work is based on), and network motif (the key component we study).

Customer Value Prediction. Customer value is regarded as one of the most reliable indicators in Customer Relationship Management (CRM) for measuring the profitability of the customer relationship and evaluating marketing decisions [1, 9]. The focus on CRM of companies in recent years has brought higher profitability to the firms [11, 64]. Prior works on customer value prediction focus on the utilization of individual historical purchase behaviors. The historically-popular model — RFM model segments customers based on their purchase recency, frequency and monetary value [11]. Kumar et al. [33] further utilize Markov processes to model fine-grained customer historical behavior data. The prosperity of Big Data allows researchers to elaborate features and utilize more complicated models. Vanderveld et al. incorporate features quantifying customers' engagement, such as the number of email clicks [56]. Chamberlain et al. use embeddings to replace some handcrafted features [17]. Chen et al. adopt deep learning methods [19]. However, they ignore the influence of social relationships on customer value. Complementing these works, this paper leverages social networks to model the social influence among customers in customer value prediction with MAG.

Graph Neural Networks. Many machine learning tasks are involved with data generated from non-Euclidean domains and represented as graphs with a set of nodes and their complex relationships [60, 66]. Recently, GNNs have achieved the art-of-state performances in many applications, e.g., recommendation [62], traffic forecasting [65], and social influence prediction [47]. GNNs extract high-order node representations by stacking multiple GNN layers. However, in many experiments, it has been observed that deeper models could not improve the performance [30]. It is mostly because stacking multiple layers can introduce high-order information but alongside with noisy information from an exponentially large number of neighbors [66]. To the best of our knowledge, we are the first to introduce GNNs into the application of customer value prediction. More importantly, our proposed MMA and GF can effectively limit the noisy information propagated by uninformative neighbors and retain informative higher-order structures.

Network Motif. Regarded as simple building blocks of complex networks, network motifs (or graphlets) [38] play essential roles in different kinds of networks [7, 32]. Specifically, triangular motifs are proved significant in social networks [28], which can be used to measure clustering coefficients [59] and model social influence [5]; feed-forward loops serve as information processing in neurons and gene regulation networks; three chains represent energy flows in food webs [38]. Multiple works utilize motifs to capture higher-order structures in graph-based tasks. Rossi et al. propose the notion of higher-order network embeddings and give various motif-based formulations [48]. Lee et al. propose motif-based graph attention networks that allow each node to select the most relevant neighborhood [34]. Sankar et al. propose a metagraph convolution operation to extract features from local structured neighborhoods in attributed heterogeneous information networks. However, prior works lack in-depth analyses of the function of motifs in models.

To fill this gap, we conduct extensive experiments to validate and interpret the function of motifs in complex social networks. Specifically, in line with earlier observations made across different social networks [5, 28], we demonstrate that the "fully-connected" motif, representing a close social relationship, is an essential network structure. Moreover, we capture another important motif – the "down-linked mutual dyad" motif, which uncovers the interplay between social influence and economic behaviors in a newly-emerging socio-technical system, social e-commerce. Indeed, customers' social and economic lives are constantly coupled in real life [24]. For example, people always turn to their friends for advice when making purchase decisions. Recently, inspired by the interplay, many platforms, e.g., Pinduoduo, Beidian, and Pinterest, incorporate social factors into their systems and have achieved great success. We originally delve into the interplay between social and economic lives from the perspective of networks and find that the "down-linked mutual dyad" motif modulates the formation of the socio-economic networks. After careful validation across various platforms, our findings could contribute to a broader field of network science and offer implications on developing future socio-technical systems that better integrate customers' social and economic lives.

7 DISCUSSION AND CONCLUSION

Customer value is the emotional bond established between a customer and a producer after the customer has used the service and found the service provides an added value [12], which means predicting customer value is estimating the perception of what the service is worth to the customer [21], rather than merely scoring individual customers. Considering the model might introduce potential biases to customers with different demographic attributes, and in turn, might cause some ethical issues, we make an in-depth analysis of the model's performance on different demographic groups.

We find out that MAG performs the best on the subgroup of unknown genders, unknown age groups, and unknown city levels, which might be attributed to the fact that most customers do not report their genders (85%), age groups (89%), as well as city levels (66%) on the Beidian dataset, and thereby the model can predict customers in this group accurately. Future work needs to further enhance the performances by making full use of the information of the "unknown" subgroup. Besides, it is worth mentioning that the Epinions dataset does not involve customers' demographics. The superior performances indicate MAG can serve customers without the need for fine-grained personal information, which might potentially alleviate the ethical and privacy concerns of some customers.

Our proposed MAG outperforms the state-of-the-art baselines on the two real-world datasets, which are collected from Beidian and Epinions respectively. Specifically, Beidian is a social e-commerce platform, where customer value manifests in the amount of money that customers are willing to spend on the platform. While Epinions is a product review platform, where the number of product reviews that customers are willing to write on the platform indicates customer value. The superior overall performances of MAG on both datasets demonstrate that it might potentially generalize to different scenarios, even when fine-grained personal information is not available.

In conclusion, to the best of our knowledge, we are the first to incorporate social network structures into customer value prediction. We discover that network structures, especially higher-order structures, are of great help to customer value prediction. We propose the framework MAG to model the effects of customer attributes and social influence jointly. Extensive evaluations and analyses reveal the superior performances of our method over baselines. Finally, through analyzing the function of motifs in social networks, we discover that more diverse motifs cannot guarantee better performances. Besides, we demonstrate that motifs are unequally informative and the "fully-connected" and the "down-linked mutual dyad" motifs are effective in predicting customer value. These findings provide new insights into understanding socio-economic relationships among customers and identifying high-value customers.

8 ACKNOWLEDGEMENT

This work was supported in part by the National Natural Science Foundation of China under U1936217, 61971267, 61972223, 61941117, 61861136003, and Beijing Natural Science Foundation under L182038.

REFERENCES

[1] Eman Aboelhamd, Hamed Shamma, and Mohamed Saleh. 2020. *Dynamic Programming Models for Maximizing Customer Lifetime Value: An Overview*. 419–445.

[2] Aris Anagnostopoulos, Ravi Kumar, and Mohammad Mahdian. 2008. Influence and correlation in social networks. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. 7–15.

[3] Aris Anagnostopoulos, Ravi Kumar, and Mohammad Mahdian. 2008. Influence and Correlation in Social Networks. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Las Vegas, Nevada, USA) *(KDD '08)*. Association for Computing Machinery, New York, NY, USA, 7–15.

[4] Sinan Aral, Lev Muchnik, and Arun Sundararajan. 2009. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences* 106, 51 (2009), 21544–21549.

[5] Qing Bao, W. Cheung, and J. Liu. 2016. Inferring Motif-Based Diffusion Models for Social Networks. In *IJCAI*.

[6] Vladimir Batagelj and Matjaž Zaveršnik. 2003. An O(m) Algorithm for Cores Decomposition of Networks. *CoRR* cs.DS/0310049 (10 2003).

[7] Austin R. Benson, David F. Gleich, and Jure Leskovec. 2016. Higher-order organization of complex networks. *Science* 353, 6295 (2016), 163–166.

[8] Leonard Berkowitz and Richard M Lundy. 1957. Personality characteristics related to susceptibility to influence by peers or authority figures. *Journal of Personality* (1957).

[9] Robert C Blattberg, John Deighton, et al. 1996. Manage marketing by the customer equity test. *Harvard business review* 74, 4 (1996), 136.

[10] Stephen P Borgatti, Ajay Mehra, Daniel J Brass, and Giuseppe Labianca. 2009. Network analysis in the social sciences. *science* 323, 5916 (2009), 892–895.

[11] Sharad Borle, Siddharth S. Singh, and Dipak C. Jain. 2008. Customer Lifetime Value Measurement. *Management Science* 54, 1 (2008), 100–112.

[12] Howard E. Butz and Leonard D. Goodstein. 1996. Measuring customer value: Gaining the strategic advantage. *Organizational Dynamics* 24, 3 (1996), 63 – 77. https://doi.org/10.1016/S0090-2616(96)90006-6

[13] Hancheng Cao, Zhilong Chen, Mengjie Cheng, Shuling Zhao, Tao Wang, and Yong Li. 2020. You Recommend, I Buy: How and Why People Engage in Instant Messaging Based Social Commerce. *arXiv preprint arXiv:2011.00191* (2020).

[14] Hancheng Cao, Zhilong Chen, Fengli Xu, Tao Wang, Yujian Xu, Lianglun Zhang, and Yong Li. 2020. When Your Friends Become Sellers: An Empirical Study of Social Commerce Site Beidian. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 14. 83–94.

[15] Herbert Castéran, Lars Meyer-Waarden, and Werner Reinartz. 2017. *Modeling Customer Lifetime Value, Retention, and Churn*. Springer International Publishing, Cham, 1–33.

[16] Damon Centola. 2018. *How Behavior Spreads: The Science of Complex Contagions*.

[17] Benjamin Paul Chamberlain, Ângelo Cardoso, C.H. Bryan Liu, Roberto Pagliari, and Marc Peter Deisenroth. 2017. Customer Lifetime Value Prediction Using Embeddings. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Halifax, NS, Canada) *(KDD '17)*. Association for Computing Machinery, New York, NY, USA, 1753–1762.

[18] Chih-Ming Chen, Ting-Hsiang Wang, Chuan-Ju Wang, and Ming-Feng Tsai. 2019. SMORe: Modularize Graph Embedding for Recommendation. In *Proceedings of the 13th ACM Conference on Recommender Systems (RecSys '19)*.

[19] Pei Pei Chen, Anna Guitart, Ana Fernández del Río, and Africa Periánez. 2018. Customer Lifetime Value in Video Games Using Deep Learning and Parametric Models. In *2018 IEEE International Conference on Big Data (Big Data)*. 2134–2140.

[20] Zhilong Chen, Hancheng Cao, Fengli Xu, Mengjie Cheng, Tao Wang, and Yong Li. 2020. Understanding the Role of Intermediaries in Online Social E-commerce: An Exploratory Study of Beidian. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–24.

[21] Zhan Chen and Alan J Dubinsky. 2003. A conceptual model of perceived customer value in e-commerce: A preliminary investigation. *Psychology & Marketing* 20, 4 (2003), 323–347.

[22] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555* (2014).

[23] Harris Drucker, Christopher JC Burges, Linda Kaufman, Alex J Smola, and Vladimir Vapnik. 1997. Support vector regression machines. In *Advances in neural information processing systems*. 155–161.

[24] Mark Granovetter. 1985. Economic action and social structure: The problem of embeddedness. *American journal of sociology* 91, 3 (1985), 481–510.

[25] Mark S. Granovetter. 1973. The Strength of Weak Ties. *Amer. J. Sociology* 78, 6 (1973), 1360–1380.

[26] Sunil Gupta, Dominique Hanssens, Bruce Hardie, Wiliam Kahn, V Kumar, Nathaniel Lin, Nalini Ravishanker, and S Sriram. 2006. Modeling customer lifetime value. *Journal of service research* 9, 2 (2006), 139–155.

[27] Nick Hajli, Julian Sims, Arash H Zadeh, and Marie-Odile Richard. 2017. A social commerce investigation of the role of trust in a social networking site on purchase intentions. *Journal of Business Research* 71 (2017), 133–141.

[28] Paul W. Holland and Samuel Leinhardt. 1977. A Method for Detecting Structure in Sociometric Data. In *Social Networks*, Samuel Leinhardt (Ed.). Academic Press, 411 – 432.

[29] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. 2017. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In *Proceedings of the 31st International Conference on Neural Information Processing Systems* (Long Beach, California, USA) *(NIPS'17)*. Curran Associates Inc., Red Hook, NY, USA, 3149–3157.

[30] Thomas N. Kipf and Max Welling. 2017. Semi-Supervised Classification with Graph Convolutional Networks. In *International Conference on Learning Representations (ICLR)*.

[31] Jon M Kleinberg. 1999. Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)* 46, 5 (1999), 604–632.

[32] Gueorgi Kossinets and Duncan J. Watts. 2006. Empirical Analysis of an Evolving Social Network. *Science* 311, 5757 (2006), 88–90.

[33] Vijaya Kumar. 2008. *Customer Lifetime Value: The Path to Profitability*. Now Publishers.

[34] John Boaz Lee, Ryan A. Rossi, Xiangnan Kong, Sungchul Kim, Eunyee Koh, and Anup Rao. 2019. Graph Convolutional Networks with Motif-Based Attention. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management* (Beijing, China) *(CIKM '19)*. Association for Computing Machinery, New York, NY, USA, 499–508.

[35] Guohao Li, Matthias Muller, Ali Thabet, and Bernard Ghanem. 2019. DeepGCNs: Can GCNs Go As Deep As CNNs?. In *The IEEE International Conference on Computer Vision (ICCV)*.

[36] Yunfei Lu, Linyun Yu, Peng Cui, Chengxi Zang, Renzhe Xu, Yihao Liu, Lei Li, and Wenwu Zhu. 2019. Uncovering the Co-driven Mechanism of Social and Content Links in User Churn Phenomena. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 3093–3101.

[37] Spyros Makridakis. 1993. Accuracy measures: theoretical and practical concerns. *International Journal of Forecasting* 9, 4 (1993), 527–529.

[38] Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon. 2002. Network Motifs: Simple Building Blocks of Complex Networks. *Science* 298, 5594 (2002), 824–827.

[39] Kathryn C Monahan, Laurence Steinberg, and Elizabeth Cauffman. 2009. Affiliation with antisocial peers, susceptibility to peer influence, and antisocial behavior during the transition to adulthood. *Developmental psychology* 45, 6 (2009), 1520.

[40] Federico Monti, Karl Otness, and Michael M Bronstein. 2018. Motifnet: a motif-based graph convolutional network for directed graphs. In *2018 IEEE Data Science Workshop (DSW)*. 225–228.

[41] Mark EJ Newman. 2003. The structure and function of complex networks. *SIAM review* 45, 2 (2003), 167–256.

[42] Margherita Pagani and Alessandra Mirabello. 2011. The influence of personal and social-interactive engagement in social TV web sites. *International Journal of Electronic Commerce* 16, 2 (2011), 41–68.

[43] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. 1999. *The PageRank Citation Ranking: Bringing Order to the Web.* Technical Report 1999-66. Stanford InfoLab. Previous number = SIDL-WP-1999-0120.

[44] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. 2014. Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. 701–710.

[45] Bryan Perozzi, Vivek Kulkarni, Haochen Chen, and Steven Skiena. 2017. Don't Walk, Skip! Online Learning of Multi-Scale Network Embeddings. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017* (Sydney, Australia) *(ASONAM '17)*. Association for Computing Machinery, New York, NY, USA, 258–265. https://doi.org/10.1145/3110025.3110086

[46] Robert J Prill, Pablo A Iglesias, and Andre Levchenko. 2005. Dynamic Properties of Network Motifs Contribute to Biological Network Organization. *PLOS Biology* 3, 11 (October 2005), 1–1.

[47] Jiezhong Qiu, Jian Tang, Hao Ma, Yuxiao Dong, Kuansan Wang, and Jie Tang. 2018. DeepInf: Social Influence Prediction with Deep Learning. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'18)*.

[48] Ryan A. Rossi, Nesreen K. Ahmed, and Eunyee Koh. 2018. Higher-Order Network Representation Learning. In *Companion Proceedings of the The Web Conference 2018* (Lyon, France) *(WWW '18)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 3–4.

[49] Ryan A. Rossi, Rong Zhou, and Nesreen K. Ahmed. 2018. Deep Inductive Network Representation Learning. In *Companion Proceedings of the The Web Conference 2018* (Lyon, France) *(WWW '18)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 953–960.

[50] J. Dan Rothwell. 1999. *In the Company of Others: An Introduction to Communication.* McGraw-Hill Higher Education.

[51] Aravind Sankar, Xinyang Zhang, and Kevin Chen-Chuan Chang. 2019. Meta-GNN: Metagraph Neural Network for Semi-supervised learning in Attributed Heterogeneous Information Networks. *ASONAM. IEEE* (2019), 137–144.

[52] Jiliang Tang, Huiji Gao, and Huan Liu. 2012. mTrust: Discerning multi-faceted trust in a connected world. In *Proceedings of the fifth ACM International Conference on Web Search and Data Mining*. ACM, 93–102.

[53] Jiliang Tang, Huiji Gao, Huan Liu, and Atish Das Sarma. 2012. eTrust: Understanding trust evolution in an online world. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 253–261.

[54] Jian Tang, Meng Qu, Mingzhe Wang, Ming Zhang, Jun Yan, and Qiaozhu Mei. 2015. LINE: Large-Scale Information Network Embedding *(WWW '15)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 1067–1077. https://doi.org/10.1145/2736277.2741093

[55] Christophe Van den Bulte and Gary L Lilien. 2001. Medical innovation revisited: Social contagion versus marketing effort. *American journal of sociology* 106, 5 (2001), 1409–1435.

[56] Ali Vanderveld, Addhyan Pandey, Angela Han, and Rajesh Parekh. 2016. An Engagement-Based Customer Lifetime Value System for E-Commerce. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (San Francisco, California, USA) *(KDD '16)*. Association for Computing Machinery, New York, NY, USA, 293–302.

[57] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. 2018. Graph Attention Networks. In *International Conference on Learning Representations*.

[58] Peter C. Verhoef, Rajkumar Venkatesan, Leigh McAlister, Edward C. Malthouse, Manfred Krafft, and Shankar Ganesan. 2010. CRM in Data-Rich Multichannel Retailing Environments: A Review and Future Research Directions. *Journal of Interactive Marketing* 24, 2 (2010), 121 – 137. Special Issue on "Emerging Perspectives on Marketing in a Multichannel and Multimedia Retailing Enviroment".

[59] Duncan J Watts and Steven H Strogatz. 1998. Collective dynamics of 'small-world'networks. *nature* 393, 6684 (1998), 440.

[60] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and S Yu Philip. 2020. A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems* (2020), 1–21.

[61] Fengli Xu, Zhenyu Han, Jinghua Piao, and Yong Li. 2019. "I Think You'll Like It": Modelling the Online Purchase Behavior in Social E-Commerce. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 65 (Nov. 2019), 23 pages.

[62] Fengli Xu, Jianxun Lian, Zhenyu Han, Yong Li, Yujian Xu, and Xing Xie. 2019. Relation-Aware Graph Convolutional Networks for Agent-Initiated Social E-Commerce Recommendation. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. Association for Computing Machinery, New York, NY, USA, 529–538.

[63] Yang Yang, Nitesh V Chawla, and Brian Uzzi. 2019. A network's gender composition and communication pattern predict women's leadership success. *Proceedings of the National Academy of Sciences* 116, 6 (2019), 2033–2038.

[64] Youjae Yi and Hoseong Jeon. 2003. Effects of Loyalty Programs on Value Perception, Program Loyalty, and Brand Loyalty. *Journal of The Academy of Marketing Science - J ACAD MARK SCI* 31 (06 2003).

[65] Bing Yu, Haoteng Yin, and Zhanxing Zhu. 2017. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. *arXiv preprint arXiv:1709.04875* (2017).

[66] Jie Zhou, Ganqu Cui, Zhengyan Zhang, Cheng Yang, Zhiyuan Liu, Lifeng Wang, Changcheng Li, and Maosong Sun. 2018. Graph Neural Networks: A Review of Methods and Applications. arXiv:1812.08434 [cs.LG]