



HA-CMNet: A Driver CTR Model for Vehicle-Cargo Matching in O2O Platform

Zilong Jiang^{1,3}, Xiang Zuo¹, Kaifu Yuan², Lin Li^{3(✉)}, Dali Wang¹,
and Xiaohui Tao⁴

¹ Department of Computer Science and Technology, Guizhou University of Finance
and Economics, Guiyang, China

² School of Business Administration, Guizhou University of Finance and Economics,
Guiyang, China

³ School of Computer Science and Artificial Intelligence, Wuhan University of
Technology, Wuhan, China
cathylilin@whut.edu.cn

⁴ University of Southern Queensland, Toowoomba, Australia

Abstract. Vehicle-cargo matching is a key task in freight O2O platform, which involves the complex interactions of drivers, vehicles, cargos, cargo owners and environmental context. Many existing works mainly study the matching of vehicle routing problems, the matching based on the credit evaluation of both drivers and cargo owners, and the matching based on game theory from the perspective of management. Since the freight O2O platform is also the producer of big data, this study proposes a driver CTR prediction model for vehicle-cargo matching task from the perspective of data mining. Specifically, we first use the bottom-level attention network to model fine-grained preferences in driver historical behaviors, such as route interest and search interest, as well as fine-grained preferences such as vehicle type and vehicle length interest, and route interest in cargo owner historical behaviors. Then, the driver basic profile vector, cargo owner basic profile vector, cargo description vector, driver preferences vector and cargo owner preferences vector are feeded into the neural network composed of two deep components for feature interactions learning, and then a top-level attention network is used to learn the influencing factors of different information on the vehicle-cargo matching task. Finally, a multi-classifier is used for matching prediction. We conduct comprehensive experiments on real dataset, and the results show that, compared with the existing solutions, considering user preferences and adopting deep components collaborative modeling can improve the performance of vehicle-cargo matching to a certain extent, which verifies the effectiveness and superiority of the proposed model.

Keywords: Recommendation system · Online To Offline platform · Vehicle-cargo matching · User preferences modeling · Hierarchical attention network

1 Introduction

In China, individual truck drivers undertake a significant portion of the road freight volume, but these truck drivers have low loading efficiency due to asymmetric logistics and distribution information, resulting in waste of social resources such as road resources and energy consumption. In order to eliminate the pain point of information asymmetry in the industry, freight O2O, a “Internet +logistics” model, has emerged. Freight O2O is a vehicle and cargo information matching platform for road logistics, in which cargo owners release vehicle demand (including route, time, cargo type, weight and volume, etc.) on the Internet and vehicle owners choose whether or not to receive orders. This platform makes use of huge industry information resources on the Internet and integrates supply chains by connecting “empty vehicles” and “cargo sources” based on big data technology. The intermediate link is omitted and the transportation cost and logistics management cost are reduced. Typical examples include intercity freight platforms such as YMM and Huochebang, and intra-city freight platforms such as Huolala.

As a key module of freight O2O platform, vehicle-cargo matching determines its stowing efficiency. Accurate and efficient vehicle-cargo matching can quickly find the most suitable vehicle owner to serve the cargo owner, which not only improves the operation efficiency of the vehicle owner, but also saves the time and cost of the cargo owner, effectively achieving the optimal allocation of resources. However, due to the complexity of the matching process involving multiple factors such as drivers, cargos, cargo owners, road transportation, and freight rates, existing researchers in the logistics field mainly focus on route optimization [1, 2], multi-objective programming [3, 4], credit evaluation [5–7], and freight rate prediction. The industry is paying more attention to the architecture and support technology of the freight O2O platform, their research on vehicle-cargo matching mainly focuses on mechanical matching using first-level indicators such as “location+route+freight”. Therefore, the existing matching mode cannot fully meet the needs of both sides of the stow, and the matching accuracy and efficiency still need to be improved.

Even a small improvement in the accuracy of vehicle-cargo matching can effectively improve the experience of drivers and cargo owners, better connect “empty vehicles” with “cargo sources”, and improve logistics costs and transportation efficiency to a certain extent. Based on the large-scale data accumulated by freight O2O platform, this study design an intelligent matching model considering user preferences to mine the potential rules contained in the historical information of freight data. Specifically, on the basis of traditional selection of first-level indicators such as “location+route+freight”, we focus on exploring the potential interests of both drivers and cargo owners, and then adopt multi-component collaborative feature interactions learning module to obtain high-order abstract feature vectors containing complex correlation relationships, and finally feed them into the prediction module to calculate the final predicted value.

The main contributions of this paper are summarized as follows:

- In order to improve the matching efficiency between drivers and cargos on the freight O2O platform, we propose a personalized recommendation algorithm for cargos to drivers, named HA-CMNet. This model can not only mine the potential preference information of drivers and cargo owners, but also learn the complex correlation relationships between multiple features by using the fitting ability of deep components, which will help improve the accuracy of vehicle-cargo matching.
- We design a special multi-party information fusion module considering user preferences. Specifically, the bottom-level attention network is first used to model various fine-grained preferences in drivers’ historical behaviors and cargo owners’ historical behaviors, and then the top-level attention network is used to learn the influence factors of different information on vehicle -cargo matching task.
- We conduct extensive experiments on real world dataset, the results show that the proposed model achieves better results on the current task compared to the advanced and most relevant mainstream baselines. In addition, in order to verify the effectiveness of the proposed model, the influence of key parameters and model structure on performance is studied extensively and deeply.

The rest of this paper is organized as follows. In Sect. 2, the problem studied are briefly stated and the proposed model and its architecture are described in detail. In Sect. 3, the comprehensive experiments are conducted to verify the performance of the proposed model. In Sect. 4, the related work is introduced. In Sect. 5, the study work is summarized and the future work is prospected.

2 Our Approach

2.1 Problem Statement

On the premise that both driver and cargo owner are registered as members of the freight O2O platform, our study takes the driver as the reference point, and the matching module in the system pushes appropriate cargos information to the driver. This process involves the following steps:

Step 1: The driver logs in to the platform and release the information related to vehicle source. The system selects cargos from the original database that meet the first level indicators such as “origin-destination”, “vehicle type-vehicle length”, and “transportation cost” based on the driver’s target route;

Step 2: If the cargos information obtained in Step 1 is too much, the intelligent matching model will further select the accurate target cargos from the sources set satisfying Step 1 by combining multiple second-level indicators of driver, cargos, cargo owner and context, and then push them to the driver after sorting.

Step 3: If the cargos information obtained in Step 1 is too little, the intelligent matching model will select relevant target cargos from the original source database by combining the driver’s location information and a variety of second-level indicators such as driver, cargos, cargo owner, and context, and push them to the driver after sorting.

This study mainly focuses on Step 2 and Step 3, and the second-level indicators will vary according to specific scenarios.

2.2 Architecture of the Proposed HA-CMNet Model

In this section, we describe the proposed HA-CMNet model, which consists of four functional modules, namely input data representation module, multi-party information fusion module considering user preferences, feature interactions learning module and matching calculation and loss function module. The complete model is a hybrid structure composed of these four modules, as shown in Fig. 1.

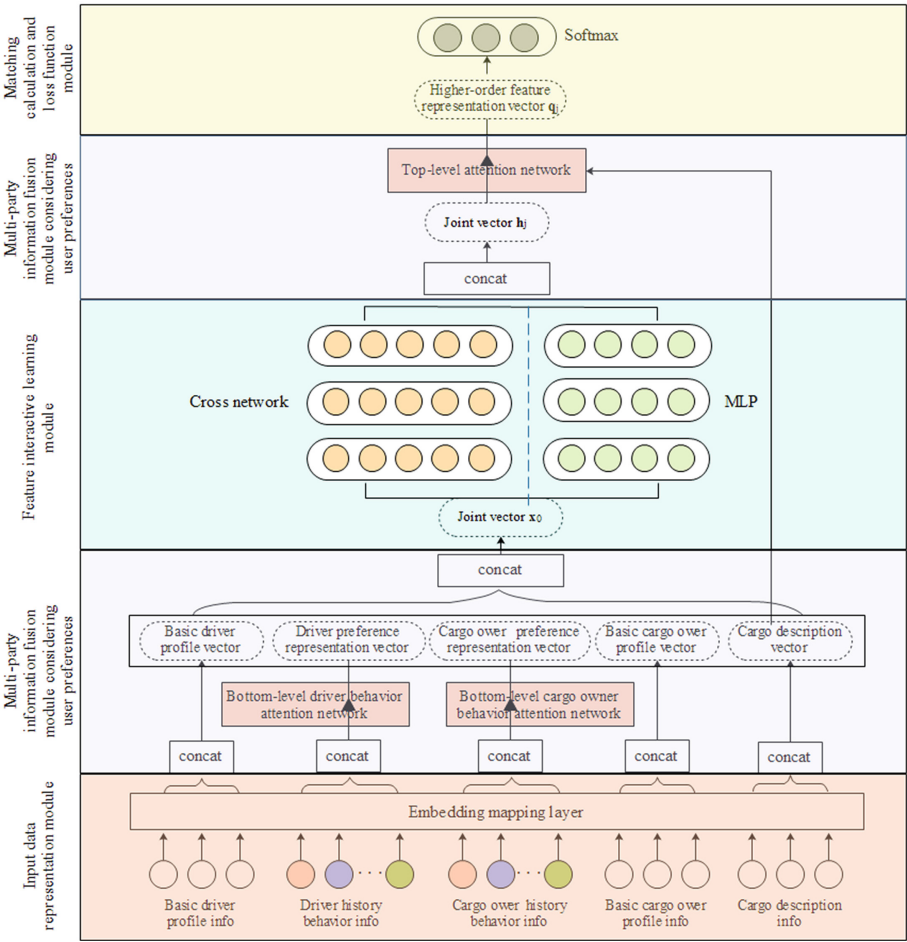


Fig. 1. Architecture of the HA-CMNet model.

The role of each module is as follows:

- (1) Input data representation module: corresponding to the light red part at the bottom of Fig. 1, it is an embedding mapping layer, which is used to convert sparse features of input data into low-dimensional dense embedding vectors. See Sect. 2.3 for a detailed description.
- (2) Multi-party information fusion module considering user preferences: corresponding to the light gray part in the middle of Fig. 1, it is a two-layer attention network, which is used to obtain the user preferences representation and the influence factors of different information on the vehicle-cargo matching task. See Sect. 2.4 for a detailed description.
- (3) Feature interactions learning module: corresponding to the light blue part in the middle of Fig. 1, it consists of a cross network [8] and multilayer perceptron (MLP), which is used to learn potential correlation relations from multi-party information and obtain higher-order feature representation. See Sect. 2.5 for a detailed description.
- (4) Matching calculation and loss function module: corresponding to the light yellow part at the top of Fig. 1, it is a Softmax classifier with three output nodes, which is used to calculate and output the final prediction results. See Sect. 2.6 for a detailed description.

The above modules and their functions are described in detail later in this section.

2.3 Input Data Representation

We sort out five types of information from the original data files: driver basic profile, driver historical behaviors, cargo owner historical behaviors, cargo owner basic profile, and cargo description. This information contains a large number of categorical field features such as the driver ID, gender, age, vehicle type, vehicle length, etc. There are search times, order number, order days, order freight and other numerical features. And the shipping routes and other text features.

As for the categorical field features, we first represent them as One-hot vectors, and then convert the One-hot vectors into low-dimensional dense embedding vectors via the embedding mapping layer. For example, the embedding mapping network of the feature “the city where the driver installs the app” can be represented by a matrix $\mathbf{V} = [\mathbf{e}_{app1}, \mathbf{e}_{app2}, \dots, \mathbf{e}_{appK}] \in R^{K \times d_v}$, where K is the total number of different cities, and $\mathbf{e}_{appj} \in R^{d_v}$ is the embedding vector with dimension d_v of item j , and the corresponding embedding vector can be obtained by looking up the table. The numerical features can be used directly after being normalized. For text features such as historical routes, we use word2vec to represent as $\mathbf{x}_b = [\mathbf{e}_{l1}, \mathbf{e}_{l2}, \dots, \mathbf{e}_{lt}, \dots, \mathbf{e}_{lT}] \in R^{T \times d_e}$, where \mathbf{e}_{lt} is the vector of the t -th route, d_e is the dimension of \mathbf{e}_{lt} , and T is the number of routes.

Finally, the complex information contained in each sample is summarized into driver basic profile vector \mathbf{x}_{dbp} , driver historical behaviors vector \mathbf{x}_{dhh} , cargo owner historical behaviors vector \mathbf{x}_{cohb} , cargo owner basic profile vector \mathbf{x}_{cobp} , and cargo description vector \mathbf{x}_{cd} after primary key association. These vectors are concatenated into the joint input vector $\mathbf{x}_{input} = [\mathbf{x}_{dbp}, \mathbf{x}_{dhh}, \mathbf{x}_{cohb}, \mathbf{x}_{cobp}, \mathbf{x}_{cd}]$ of

the sample . In addition, the class label of each sample is a One-hot vector represented by three kinds of tags: browse cargos, click cargos and make phone calls.

2.4 Multi-party Information Fusion Considering User Preferences

For driver users, whether the transaction of the target cargo is finally concluded is not only related to the first-level index of “location + route + freight rate”, but also involves the actual willingness of both the driver and the cargo owner. Understanding and restoring the potential intentions of both the driver and the cargo owner is the key issue to be addressed in this section. The attention mechanism can consider the weight relationship between different elements on the basis of a global perspective. Taking advantage of this advantage, we adopt hierarchical attention network [9] to first model the fine-grained preferences of driver and cargo owner, and then learn the influence factors of different information on the vehicle-cargo matching task, and finally incorporate them into the subsequent matching prediction calculation.

2.4.1 The Bottom-Level Attention Network Based User Preferences Modeling

- *Driver Preferences Modeling*

We design a specialized driver behavior attention network to model driver preferences, in order to obtain driver preferences representation vector. The input of this part comes from the joint input vector \mathbf{x}_{input} . First, we extract each behavior \mathbf{x}_{dhbi} from the driver historical behaviors vector $\mathbf{x}_{dhh} = [\mathbf{x}_{dhb1}, \dots, \mathbf{x}_{dhbM}]$, and then concatenate it with the driver basic profile vector \mathbf{x}_{dbp} , the cargo owner basic profile vector \mathbf{x}_{cobp} , and the cargo description vector \mathbf{x}_{cd} to obtain the input vector $[\mathbf{x}_{dbp}, \mathbf{x}_{cobp}, \mathbf{x}_{cd}, \mathbf{x}_{dhbi}]$ for this part. The output is the learned driver preferences representation vector \mathbf{x}'_{dhh} .

Firstly, we model the attention score α_m for each driver behavior, as shown in Eq. (1),

$$\alpha_m = \mathbf{w}_1^T \cdot s(\mathbf{W}_1[\mathbf{x}_{dbp}, \mathbf{x}_{cobp}, \mathbf{x}_{cd}, \mathbf{x}_{dhbi}]), \tag{1}$$

where $\mathbf{w}_1 \in R^{d_2}$ and $\mathbf{W}_1 \in R^{d_2 \times d_1}$ are parameters of the driver behavior attention network, d_1 is the dimension of the input vector of the attention network, d_2 is the dimension of the output vector of the attention network, and $s(x)$ is the nonlinear activation function.

Then, by normalizing the attention score α_m in Eq. (1), we obtain the final attention score α'_m as follows,

$$\alpha'_m = \frac{\exp(\alpha_m)}{\sum_{l=1}^M \exp(\alpha_l)}, \tag{2}$$

Finally, we concatenate the driver behavior vectors weighted by the final attention scores into a driver preferences representation vector as follows,

$$\mathbf{x}'_{dhb} = \sum_{m=1}^M \alpha'_m \mathbf{x}_{dhbm}. \tag{3}$$

• *Cargo Owner Preferences Modeling*

Similarly, we design a specialized cargo owner behavior attention network to model cargo owner preferences, in order to obtain cargo owner preferences representation vector. The input of this part comes from the joint input vector \mathbf{x}_{input} . First, we extract each behavior \mathbf{x}_{cohbi} from the cargo owner historical behaviors vector $\mathbf{x}_{cohb} = [\mathbf{x}_{cohb1}, \dots, \mathbf{x}_{cohbN}]$, and then concatenate it with the driver basic profile vector \mathbf{x}_{dbp} , the cargo owner basic profile vector \mathbf{x}_{cobp} , and the cargo description vector \mathbf{x}_{cd} to obtain the input vector $[\mathbf{x}_{dbp}, \mathbf{x}_{cobp}, \mathbf{x}_{cd}, \mathbf{x}_{cohbi}]$ for this part. The output is the learned cargo owner preferences representation vector \mathbf{x}'_{cohb} .

Firstly, we model the attention score β_n for each cargo owner behavior, as shown in Eq. (4),

$$\beta_n = \mathbf{w}_2^T \cdot s(\mathbf{W}_2[\mathbf{x}_{dbp}, \mathbf{x}_{cobp}, \mathbf{x}_{cd}, \mathbf{x}_{cohbi}]), \tag{4}$$

where $\mathbf{w}_2 \in R^{d_3}$ and $\mathbf{W}_2 \in R^{d_4 \times d_3}$ are parameters of the cargo owner behavior attention network, d_3 is the dimension of the input vector of the attention network, d_4 is the dimension of the output vector of the attention network, and $s(x)$ is the nonlinear activation function.

Then, by normalizing the attention score β_n in Eq. (4), we obtain the final attention score β'_n as follows,

$$\beta'_n = \frac{\exp(\beta_n)}{\sum_{l=1}^M \exp(\beta_l)}, \tag{5}$$

Finally, we concatenate the cargo owner behavior vectors weighted by the final attention scores into a cargo owner preferences representation vector as follows,

$$\mathbf{x}'_{cohb} = \sum_{n=1}^N \beta'_n \mathbf{x}_{cohbn}. \tag{6}$$

2.4.2 The Top-Level Attention Network Based Influence Modeling of Different Information

Information from different sources has different activity distribution and is of different importance to vehicle-cargo matching. We adopt the top-level attention network to model the influence of different information on the final matching task. The input of this part is six kinds of information, including driver basic profile vector \mathbf{x}_{dbp} , cargo owner basic profile vector \mathbf{x}_{cobp} , cargo description vector \mathbf{x}_{cd} , driver preferences representation vector \mathbf{x}'_{dhb} and cargo owner preferences

representation vector \mathbf{x}'_{cohb} obtained in Sect. 2.4.1, and higher-order abstract feature representation \mathbf{h}_j obtained in the deep components of subsequent Sect. 2.5. The output is the corresponding influence factor of various information and the higher-order feature representation \mathbf{q}_j .

Firstly, we model the attention score $\delta_{jj'}$ for each kind of information, as shown in Eq. (7),

$$\delta_{jj'} = \mathbf{w}^T \cdot \tanh(\mathbf{V}\mathbf{s}_{j'} + \mathbf{W}\mathbf{h}_j), \tag{7}$$

where \mathbf{w} , \mathbf{V} and \mathbf{W} are parameters of the top-level attention network, $\mathbf{s}_{j'}$ ($j' = 1, 2, 3, 4, 5$) is the vector representations of five kinds of information obtained from the bottom-level attention network. Specifically, $\mathbf{s}_1 = \mathbf{x}_{dbp}$, $\mathbf{s}_2 = \mathbf{x}_{cobp}$, $\mathbf{s}_3 = \mathbf{x}_{cd}$, $\mathbf{s}_4 = \mathbf{x}'_{dhhb}$, $\mathbf{s}_5 = \mathbf{x}'_{cohb}$. And \mathbf{h}_j is the high-level feature representation obtained by subsequent feature interactions learning module when predicting the j -th classification.

Then, by normalizing the attention score $\delta_{jj'}$ in Eq. (7), we can obtain the final attention score $\delta'_{jj'}$ for predicting the above five aspects of information in multiple classifications, which is the impact factor of different information on the vehicle-cargo matching task. The calculation formula is as follows,

$$\delta'_{jj'} = \frac{\exp(\delta_{jj'})}{\sum_{l=5}^M \exp(\delta_{jl})}, \tag{8}$$

Finally, when predicting the j -th class tag, the higher-order feature representation vector \mathbf{q}_j obtained from the top-level attention network can be calculated as:

$$\mathbf{q}_j = \sum_{j'=1}^5 \delta'_{jj'} \cdot \mathbf{s}_{j'}. \tag{9}$$

2.5 Feature Interactions Learning

In order to fully explore the potential correlation relationships between multiple features, we use cross network and MLP for collaborative modeling, and learn higher-order abstract feature representation in both explicit and implicit ways. The input of this part is a joint vector $\mathbf{x}_0 = [\mathbf{x}_{dbp}, \mathbf{x}_{cobp}, \mathbf{x}_{cd}, \mathbf{x}'_{dhhb}, \mathbf{x}'_{cohb}]$ concatenated by the driver basic profile vector \mathbf{x}_{dbp} , the cargo basic profile vector \mathbf{x}_{cobp} , the cargo description vector \mathbf{x}_{cd} obtained in Sect. 2.3, and the driver preferences representation vector \mathbf{x}'_{dhhb} and the cargo owner preferences representation vector \mathbf{x}'_{cohb} obtained in Sect. 2.4.1. The output is a joint vector \mathbf{h}_j concatenated by the higher-order feature representations obtained from the cross network and MLP.

- *Cross Network based Explicit Feature Interactions Learning*

We use a cross network [8] to explicitly learn vector-level higher-order cross features, and the input of cross network is the joint vector \mathbf{x}_0 mentioned above. The cross network is composed of cross layers, with each layer having the following Eq. (10):

$$\mathbf{x}_{l+1} = \mathbf{x}_0 \mathbf{x}_l^T \mathbf{w}_l + \mathbf{b}_l + \mathbf{x}_l = f(\mathbf{x}_l, \mathbf{w}_l, \mathbf{b}_l) + \mathbf{x}_l, \tag{10}$$

where $\mathbf{x}_l, \mathbf{x}_{l+1} \in R^d$ are column vectors denoting the outputs from the l -th and $(l + 1)$ -th cross layers, respectively, $\mathbf{w}_l, \mathbf{b}_l \in R^d$ are the weight and bias parameters of the l -th layer, and $l \in [1, \dots, L_1]$. Each cross layer adds back its input after a feature crossing f , and the mapping function $f : R^d \mapsto R^d$ fits the residual of $\mathbf{x}_{l+1} - \mathbf{x}_l$. A visualization of one cross layer is shown in Fig. 2.

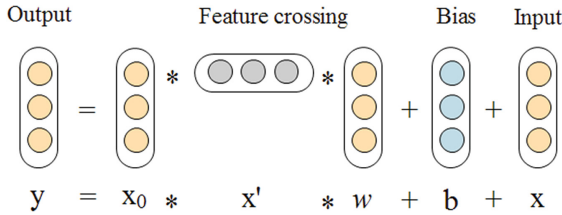


Fig. 2. Visualization of a cross layer.

The output \mathbf{x}_{L_1} of the cross network is the feature representation vector obtained from its last layer, which will be concatenated with the higher-order feature representation vector obtained from the last layer of the MLP network to form a joint vector \mathbf{h}_j .

• *MLP based Implicit Feature Interactions Learning*

We use MLP network to implicitly learn bit-level high-order cross features. The input of MLP is also the joint vector \mathbf{x}_0 mentioned above.

The forward propagation process of the network can be formally described as:

$$\mathbf{x}^l = \sigma (\mathbf{W}^l \mathbf{x}_0 + \mathbf{b}^l), \tag{11}$$

$$\mathbf{x}^{l+1} = \sigma (\mathbf{W}^{l+1} \mathbf{x}^l + \mathbf{b}^{l+1}), \tag{12}$$

where $l \in [1, \dots, L_2]$ is the number of hidden layer, \mathbf{W}^l is the connection weight matrix of the $(l - 1)$ -th and l -th layers of MLP network, \mathbf{b}^l is the offset of the l -th layer of the MLP network, σ is the nonlinear activation function, and \mathbf{x}^l is the output vector of the l -th layer.

The output \mathbf{x}_{L_2} of the MLP network is the feature representation vector obtained from its last layer, which will be concatenated with the higher-order feature representation vector obtained from the last layer of the crossover network to form a joint vector \mathbf{h}_j .

2.6 Matching Calculation and Loss Function

The task studied in this paper belongs to the multi-classification problem. We use the Softmax classifier for matching prediction calculation. The input of this part is the higher-order feature representation vector \mathbf{q}_j obtained from the top-level

attention network in Sect. 2.4.2, and the output is a predicted value calculated by the following Eq. (13),

$$\hat{y} = P(y = j) = \text{Softmax}(\mathbf{q}_j) = \frac{\exp(\mathbf{q}_j)}{\sum_{j'=0}^2 \exp(\mathbf{q}_{j'})}, j \in \{0, 1, 2\}, \quad (13)$$

In addition, we use the cross entropy loss function with regularization term to optimize the model parameters,

$$L = P(y = j) = -\frac{1}{N} \sum_{i=1}^N \left(\sum_{j=0}^2 y_i \log(\hat{y}_i) \right) + \lambda \sum_l \|w_l\|^2, \quad (14)$$

where \hat{y}_i is the predicted value calculated according to Eq. (13), y_i is the label value, N is the total number of samples, and λ is the penalty factor for the regularization term.

3 Experiments

In this section, we introduce the experiments in detail, including the dataset, evaluation metrics, baselines, comparative study, hyperparameters setting and sensitivity analysis, and ablation study.

3.1 Dataset

This study use a dataset provided by China’s largest intercity freight O2O platform, Full Truck Alliance Co. Ltd [10]. We sort out five types of information from the original file of the dataset by means of primary key association: driver basic profile including driver ID, gender, age, registration time, vehicle type, vehicle length, APP platform, device brand, etc.; driver historical behaviors including the cargos sent by the driver, recent search routes in different time periods, number of browsing sources, number of clicks on sources, and number of phone calls made; cargo owner historical behaviors including their historical shipping routes, categories, total recent shipments in different time periods, number of days shipped, the quantity of goods complained, etc.; cargo owner basic profile including ID, gender, age, authentication time, activation date, APP platform, device brand, package type, and other features; cargo description including features such as cargo ID, cargo owner ID, origin, destination, weight, volume, type, freight rate, required vehicle type, required vehicle length, etc.

3.2 Evaluation Metrics

The prediction task of this study belongs to the multi-classification problem, and we use three evaluation metrics, macro-precision, macro-recall and macro-F1 score, for evaluation [11]. Among them, macro F1-score is a balanced comprehensive index. The larger the value of these metrics, the better the classification effect.

3.3 Baselines

In order to fully verify the effectiveness of the proposed model, we compare it with several state of the art intelligent vehicle-cargo matching methods. The details of the baselines are as follows:

LightGBM [12]: This method can well model the feature interactions and is widely used in industry.

XGBoost [13]: This method can also better model the potential interactions between different features in vehicle and cargo information, and obtain valuable and interpretable high-order features, and has a good effect in industry.

DNN based competition plan [14]: This method can fit the potential law among the features, and currently some leading enterprises begin to exploit the deep learning technology for intelligent matching of vehicle, cargo and context information.

A-SENet [15]: This method introduces SENet and attention on the basis of the DNN based competition plan, which is currently a relatively novel solution.

3.4 Comparative Study

For HA-CMNet trained from scratch, we use Gaussian distribution (mean value is 0, standard deviation is 0.01) to randomly initialize model parameters. The training loss of the proposed model on the dataset is shown in Fig. 3, and the hyperparameters setting is shown in Sect. 3.5. In this section, we compare HA-CMNet with the baselines on the same task.

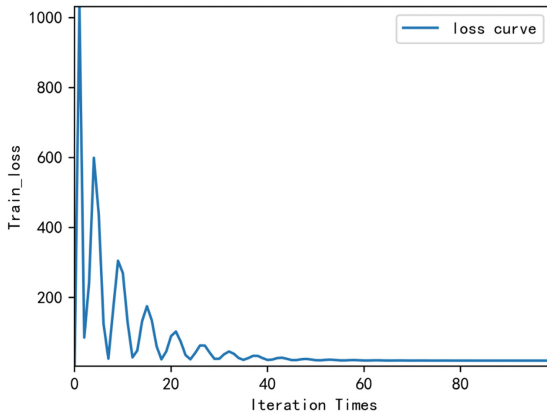


Fig. 3. Training loss of HA-CMNet on this dataset.

• Parameters Setting of Baselines

In a consistent experimental environment, in order to make baselines reach their best states, we also optimize them through many experiments. Finally, for LightGBM, we set learning_rate to 0.05, min_child_sample to 18, and max_depth

to 5. For XGBoost, we set booster to gbtrees, max_depth to 3, learning rate to 0.001, objective to multi:softprob, eval_metric to mlogloss, early_stop is 15. For the DNN based competition plan and A-SENet, we adopt an MLP structure of 100-100-100, a learning rate of 0.0001, and an optimizer of Adam. To avoid overfitting, L2 regularization is adopted and the dropout is set at 0.5.

- *Results and Comparative Analysis*

Table 1. Performance comparison of all models.

Model	Macro-precision	Macro-recall	Macro-F1 score
LightGBM	0.8104	0.7712	0.7834
XGBoost	0.7936	0.8074	0.8011
DNN based competition plan	0.8272	0.8322	0.8273
A-SENet	0.8372	0.8410	0.8351
<u>HA-CMNet</u>	<u>0.8593</u>	<u>0.8465</u>	<u>0.8502</u>

Table 1 shows the experimental results of all models on the current prediction task. LightGBM and XGBoost are traditional methods used in the industry, which train multiple decision trees to fit the association relationships in multi-party information, and mining the potential association relationships between features is proven to be helpful for prediction. Compared with XGBoost, LightGBM has faster training speed and lower computational overhead, but we find that its accuracy is slightly lower than XGBoost, which may be due to its decision tree node splitting method and tree growth method.

DNN based competition plan achieve better results than LightGBM and XGBoost, demonstrating that the DNN based solution outperforms the traditional tree based model due to its superior extraction ability of complex association relationships compared to the tree based model. Compared with DNN based competition plan, A-SENet adds SENet modules to the left and right towers, which can dynamically represent the importance of different features and improve the prediction accuracy to a certain extent.

And our proposed HA-CMNet achieve better results than DNN based competition plan and A-SENet, mainly due to the following two advantages: (i) By introducing the hierarchical attention network, selecting significant latent information from different levels is helpful to improve the modeling ability of the model; (ii) By combining cross network and MLP to learn multi-party feature interactions, prediction accuracy can be improved more effectively than separately learning the driver latent vector and the cargo owner latent vector.

In conclusion, by sufficient comparison with the mainstream baselines, the proposed model can achieve better results on the experimental dataset, which proves the validity of the HA-CMNet model design.

3.5 Hyperparameters Setting and Sensitivity Analysis

In this section, we mainly study the influence of hyperparameters setting on the performance of HA-CMNet, including: (i) MLP network structure; (ii) Number of cross network layers; (iii) Learning rate; (iv) Activation function; (v) Penalty factor; (vi) Optimizer. We adopt control variable method to conduct experiments, that is, first fix the values of other hyperparameters, and then change the value of a hyperparameter to be understood within a certain range to observe its influence on the prediction results.

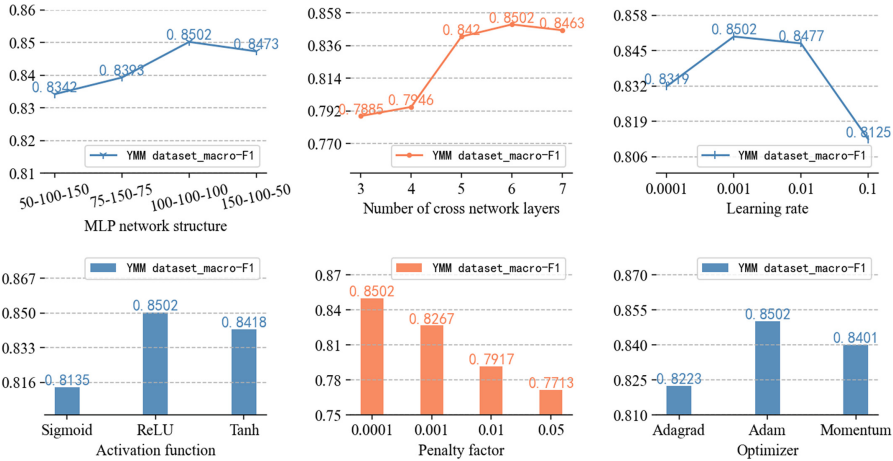


Fig. 4. Hyperparameters setting and sensitivity analysis.

As shown in Fig. 4, we test the influence of four network structures [Diamond, Constant, Incrementing, and Decreasing] on the final performance. It can be observed that when the MLP network structure is Constant, the performance of HA-CMNet is relatively best, which may be because this structure is more suitable for the current prediction task. We also find that if the number of hidden layers continues to increase, the model performance will not be significantly improved, and this phenomenon is caused by excessive parameters leading to overfitting.

For the cross network, we find that the effect is relatively stable when the number of cross layers is 6, and increasing the number of cross layers no longer improves the effect. This indicates that introducing a higher degree of feature interactions is not helpful. Finally, we choose the MLP network structure as 100-100-100, and the number of layers of the cross network is 6.

As for the learning rate, we test in four values [0.0001, 0.001, 0.01, 0.1], and find that 0.001 is better in convergence speed and accuracy. By comparing three activation functions: Sigmoid, ReLU and Tanh, it is found that ReLU is slightly

better than the other two. This may be because ReLU is a left saturation function, which alleviates the problem of gradient disappearance of neural network to a certain extent, and can accelerate the convergence rate of gradient decline.

Based on experience, the penalty factor for the regularization term is tested among four values: [0.0001, 0.001, 0.01, 0.05], and it is found that 0.0001 is slightly better. The effect is relatively good when the optimizer is Adam. The dropout rate is set to 0.5, and we find that this value matches the ReLU activation function in the experiment.

3.6 Ablation Study

In the ablation test, in order to further evaluate the design rationality and validity of the proposed model, we will analyze the influence of different components in HA-CMNet in detail.

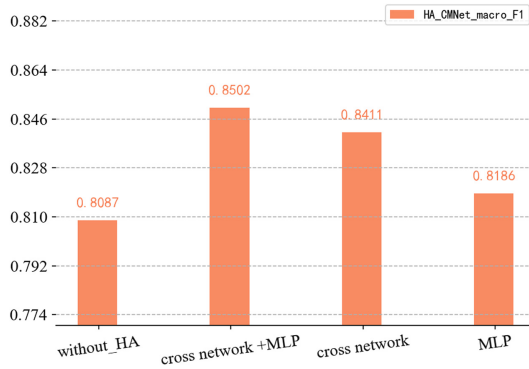


Fig. 5. The effect of the proposed models with different components.

- *The Influence of Hierarchical Attention Network on Prediction*

We test the effect of models with and without hierarchical attention networks, respectively. From Fig. 5, we can see that the former achieve significantly better results, which can be explained by the fact that it contains two levels of improvement: first, the low-level attention network focuses on fine-grained user preferences; second, the top-level attention network strengthens the influence of different information on matching prediction. This is attributed to the fact that the attention network can effectively select significant latent information from the feature level, which helps improve the proposed model’s expressive ability.

- *The Influence of Cross Network and MLP on Prediction*

We test the effect of feature interactions learning modules using three different methods: cross network+MLP, cross network, and MLP. It can be observed from Fig. 5 that the effect of MLP is relatively low, while that of cross network is significantly higher, and cross network +MLP achieves the best effect. We also find that combinations including cross network can achieve better results,

which is an interesting phenomenon that explicit modeling of high-order feature interactions has a better effect on matching prediction, and can be regarded as an experience or rule.

In conclusion, theoretically and practically, the proposed model can effectively select significant latent information by using hierarchical attention network, and can simultaneously learn feature interactions from different perspectives by using cross network and MLP, which is a reasonable and effective design mode for vehicle-cargo matching task.

4 Related Work

As a new supply and demand stowing mode, vehicle-cargo matching in the freight O2O platform attracts researchers and engineers at home and abroad to study from different perspectives due to its complexity and importance in the current logistics transportation.

Li et al. studied the one vehicle multi-point delivery service [1] in the freight O2O platform and the supply-demand matching and route planning of the zero-carload cargo business [2]. Some researchers also use intelligent algorithms such as semantic web and evolutionary algorithm to study the vehicle-cargo matching problem. For example, Gu constructed an ontology in the field of vehicle-cargo matching, studied the matching rate of road freight vehicles and cargos, and implemented the basic intelligent reasoning function of the highway freight information platform and semantic based matching between vehicles and cargos [3]. Liu constructed an information index system and vehicle-cargo matching model based on multi-objective programming, proposed a solving algorithm for vehicle-cargo matching model based on genetic algorithm, and optimized and analyzed the vehicle-cargo matching problem of Chuanhua Logistics Highway Port [4].

Some researchers also use game theory and credit evaluation methods to study problems related to vehicle-cargo matching. For example, Jia et al. constructed a bilateral user transaction game model for the vehicle-cargo matching platform, exploring the control problem of platform users evolving from multi attribution to single attribution [5]. Shao et al. selected four platforms such as huochebang, YMM as the analysis object, first constructed the competitiveness evaluation index system for the vehicle-cargo matching platform, and then used analytic hierarchy process to determine the weights of each evaluation index, finally, used the fuzzy comprehensive evaluation method to evaluate the selected platform and obtained its competitiveness level [6]. Bing used analytic hierarchy process and fuzzy comprehensive evaluation to construct a credit evaluation system for vehicle owners and cargo owners, and established a one-to-one vehicle-cargo matching scheduling model and a one-to-many vehicle-cargo matching scheduling model with the goal of minimizing matching costs [7].

The industry also begins to explore solutions based on big data technology. They adopted the embedding mapping layer and DNN to model driver latent vector and cargo latent vector, and then adopted softmaxclassifier for classifi-

cation [14]. On this basis, Fang et al. proposed a driver CTR model A-SENet, which differed in that SENet was used to calculate the cargo latent vector, while attention and SENet were used to calculate the driver latent vector [15].

5 Conclusion and Future Work

This paper proposes a driver CTR model named HA-CMNet, which aims to improve the efficiency and accuracy of vehicle-cargo matching task in freight O2O platform. The model has the following advantages: (i) It can effectively mine the preferences information of drivers and cargo owners; (ii) It can learn the influence factors of different information on the vehicle-cargo matching task; (iii) It can learn the higher-order feature interactions at both the bit-level and the vector-level, and obtain the cross features containing rich semantic information. Detailed and comprehensive experiments are conducted on the real dataset provided by Full Truck Alliance Co. Ltd., and the results show that the proposed model can achieve a certain degree of better performance than the baselines.

There are two directions for future work. Firstly, we explore the introduction of graph comparison learning [16], graph convolutional networks [17], and other graph methods [18, 19] to improve item representation accuracy; Secondly, we consider extending our solution to the order-rider matching problem in O2O platforms for life services.

Acknowledgement. This work is supported by National Natural Science Foundation of China (NO.72261003, 62276196, 62106070), Guizhou Provincial Science and Technology Project (NO.Qiankehejichu-ZK[2022]yiban019), and Key Research and Development Project of Hubei Province (NO.2021BAA030).

References

1. Li, J., Zheng, Y., Dai, B., Yu, J.: Implications of matching and pricing strategies for multiple delivery-points service in a freight O2O platform. *Transp. Res. Part E Logistics Transp. Rev.* **136**, 101871 (2020)
2. Li, J., Zhou, T., Xu, L., Dai, B.: Matching optimization decision of city LTL carpool based on time windows on the freight O2O platform. *Syst. Eng. – Theor. Pract.* **40**(4), 979–988 (2020)
3. Gu, J.: *Vehicle-Cargo Matching System Based on Semantic Web Technology*. Tsinghua University, Beijing (2013)
4. Liu, X.: *Research on the Optimization of the Matching of Truck and Cargo in the Highway Port of Transfar*. Tsinghua University, Beijing (2013)
5. Jia, X., Hai, F., Dong, R.: Control design for promoting single-homing user ratio of vehicles and cargos matching two-sided platform. *Comput. Integr. Manuf. Syst.* **23**(4), 903–912 (2017)
6. Shao, Y., Yang, M., Wang, Y.-J., Wei, Y.-Z.: Research on competitiveness of vehicle-freight matching platform for carrier. In: *15th International Conference on Service Systems and Service Management*. Hangzhou, China (2018)

7. Bing, C.: Research on Vehicle-Cargo Matching in Highway Trunk Freight Transportation - Taking T Platform as an Example. Zhejiang Sci-Tech University, Hangzhou (2018)
8. Wang, R., Fu, B., Fu, G., Wang, M.: Deep & Cross network for ad click predictions. In: 23th ACM SIGKDD Conference on Knowledge Discovery and Data Mining Workshop-ADKDD&TargetAD. Halifax, NS, Canada (2017)
9. Sun, J., Zhu, M., Jiang, Y., Liu, Y.-Z., Wu, L.: Hierarchical attention model for personalized tag recommendation. *J. Am. Soc. Inf. Sci.* **72**(2), 173–189 (2021)
10. YMM-TECH algorithm contest dataset. <https://ymmtech.ymm56.com>. Accessed 6 Feb 2020
11. Zhou, Z.-H.: Machine Learning, 2nd edn. Tsinghua University Press, Beijing (2021)
12. Ke, G., et al.: LightGBM: a highly efficient gradient boosting decision tree. In: The Thirty-First Annual Conference on Neural Information Processing Systems. Long Beach, CA, USA, pp. 3146–3154 (2017)
13. Chen, T., Guestrin, C.: XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, CA, USA, pp. 785–794 (2016)
14. DNN based competition plan. <http://api.ymm56.com/article4153.html>. Accessed 12 Oct 2022
15. Fang, F., Wang, C.: Driver click-through rate prediction model based on SENet two-tower model considering attention mechanism in vehicle-cargo matching. *Logistics Sci. Technol.* **2**, 91–97 (2022)
16. Song, X., Li, J., Lei, Q., Zhao, W., Chen, Y., Mian, A.: Bi-CLKT: Bi-graph contrastive learning based knowledge tracing. *Knowl.-Based Syst.* **241**, 108274 (2022)
17. Song, X., Li, J., Tang, Y., Zhao, T., Chen, Y., Guan, Z.: A joint graph convolutional network based deep knowledge tracing. *Inf. Sci.* **580**, 510–523 (2021)
18. Xue, G., Zhong, M., Li, J., Chen, J., Zhai, C., Kong, R.: Dynamic network embedding survey. *Neurocomputing* **472**, 212–223 (2022)
19. Cai, T., Li, J., Mian, A., Li, R.-H., Sellis, T., Yu, J.X.: Target-aware holistic influence maximization in spatial social networks. *IEEE Trans. Knowl. Data Eng.* **34**(4), 1993–2007 (2022)