Hindawi
Mobile Information Systems
Volume 2022, Article ID 2364154, 11 pages
https://doi.org/10.1155/2022/2364154

Research Article

Applying Deep Learning-Based Personalized Item Recommendation for Mobile Service in Retailor Industry

Minghua Xiao,[1] Qing Zhou,[1] Lei Lu,[1] Xingzhen Tao,[1] Wenting He,[1] and Youmei Zhou[2]

[1]*School of Information Engineering, Jiangxi College of Applied Technology, Ganzhou, China*
[2]*College of Architecture and Urban Planning, Tongji University, Shanghai, China*

Correspondence should be addressed to Youmei Zhou; 20310231@tongji.edu.cn

Received 12 April 2022; Revised 5 May 2022; Accepted 23 May 2022; Published 8 June 2022

Academic Editor: Yan Huang

Various kinds of mobile services allow integrating terminal customers as important coproducers into the whole retailer's business processes. People have enjoyed increasing popularity in the past years since they allow saving costs and increasing satisfaction. However, in some retail settings, as the technology relies on retailers providing terminals, it does not yet fully utilize the possibilities provided by mobile service, which until recently have mostly served as shopping aids. Recommendation systems can provide accurate recommendation services to users, especially in the field of e-commerce. In this study, a mobile retail terminal, Kkbox, leverages deep learning-based recommendation and self-service technologies to provide an express and personalized self-checkout retail environment without the engagement of storekeepers and cashiers. An attention-based mechanism for product personalization recommendation model is adopted, and it models the intrinsic relationship between users' historical interactions with products through a multilayer self-attentive network and then feeds the output of the multilayer self-attentive network into a GRU network with attention scores to model the evolution of users' interests. We analyze the performance of the product recommendation module based on user data from multiple perspectives, such as purchase frequency, purchase time, and product category. In the comparison experiments with some traditional recommendation methods, the recommendation accuracy of the model used in this study achieves better results. Besides, it significantly reduces the labor cost and provides enough flexibility. The time performance of app users is independent of store rush. The time for a transaction is significantly lower for app users than the regular shoppers during peak periods. The Kkbox has been deployed in several communities in Taizhou, China, to provide fast and convenient mobile retail services to residents.

1. Introduction

The retail sector has been playing a dominant role in economic activities, and it is a strategic industry of the national economy in China. China's total volume of retail sales increased from 9.4 trillion in 2012 to 21 trillion in 2016 [1]. The supermarket dominates the retail sector [2], and the chain operation development mode is widely adopted. In such a manner, stock can be replenished in large amounts, thus lowering the purchase cost of commodities shared by a number of chain stores in the same areas. This can also dynamically maintain the demand balance of each store [3]. The customers are able to buy most of the commodities needed in their daily life in the supermarket. Meanwhile, the large-scale and efficient management policies for supermarkets greatly reduce the price of commodities.

However, the operational efficiency of the supermarket is not as good as expected. There usually exists some frustrating experience when the customers go shopping in the supermarket, for example, the customer usually suffers from the long queue at the checkout counter, and with the investigation by researchers, such a long queue may be caused by the bar code scanning, which is time consuming [4–7]. Before completing the checkout process, all commodities need to be taken out from the shopping bag and scanned on the counter one by one. Such a process is also boring and time consuming; The hundreds of commodities always make customers get lost, whereas appropriate update news of commodities could directly reach their potential customers and promote the purchase. But, actually, there are very limited channels (e.g., news subscription) available. Therefore, more convenient guides and assistances are expected (e.g., apps) [8–10].

Nowadays, self-service technologies (SSTs), such as self-checkout terminals, enable customers to scan and pay for groceries without interaction with the store personal [11]. SST has been proven to be an effective way to increase consumer satisfaction and convenience at checkout by avoiding the long queues associated with traditional checkout methods. [1, 2] However, SST using barcode scanning and self-service payment devices is mainly used in large shopping centers [12] and has a limited reach. In these cases, the shopping basket is transferred to a self-checkout terminal, where the actual payment and delivery of the merchandise occurs, but in practical terms, the checkout capacity remains limited during peak periods [1]. Current self-checkout solutions can be very appropriate for the digital transformation of general grocery stores, and they are already well adopted in these stores [12]. Still, they are not applicable to the convenience store environment. Unlike regular grocery stores, customers in convenience stores usually purchase very few items. In China, convenience stores are typically located in large communities with limited space and expensive rents [13, 14]. The in-store checkout load is hefty during morning and evening rush hours.

Our study combines SSTs, mobile applications, and deep learning into a solution that provides digital innovative customer experiences in the kiosk. The contributions in this work can be summarized as follows:

(i) Our solutions are driven by unique personalized content. The technology in this article is delivered via a large display in a smart retail installation and is set up to ensure that the content perfectly conforms to the user's preferences and specific interactive application.

(ii) The system collects relevant content about customer interactions and provides measurable statistics on customer engagement.

(iii) This solution can provide promoted products based on recommendations under deep learning.

The rest of the study is organised as follows: Section 2 presents the relevant work and introduces the background of our research context, questions, and methodology. Section 3 illustrates the current situation, and based on this, we design a target shopping process adapted to our research context and provide an overview of our system architecture. Section 4 discusses the results from our field study. Section 5 concludes.

2. Related Work

2.1. Self-Service Technologies. Self-service technology (SST) is a technical interface that allows customers to coproduce services without employee interaction [11]. Retailers mainly provide SST to reduce costs and improve customer experience [2]. The most commonly cited advantages of SST are convenience and speed [1]. The users' negative experience in the process of use may mainly include the experience of forced use of self-checkout, the closure of self-checkout terminals at certain times of the day (such as night), and the fact that these terminals happen to be slow when queuing [1]. Studies have shown that SST customers tend to use self-checkout terminals to check out smaller baskets and may avoid items (such as fruits and vegetables) that require additional steps in the checkout process [15, 16]. Purchasing fewer items is a crucial reason for self-checkout, as there are only long lines at the top of the shopping basket on the traditional POS machines [17]. The service quality of SST is mainly determined by function, enjoyment, design, guarantee, and convenience. SST also positively impacts customer loyalty through customer satisfaction [2]. According to the existing literature, the current technology does not provide a self-checkout solution for the grocery retail industry [18–20].

Self-checkout is a typical self-service application. Self-checkout allows retailers to order customers to scan, pack, and pay for items they want without hiring staff. Considering that retail stores alone make more than 60 billion transactions a year, 68% of which are grocery stores, gas stores, and convenience stores. This will result in significant cost savings, as self-checkout eliminates the need to hire more staff. In addition, it can also reduce the time that customers spend waiting in line, which is one of their biggest complaints. NCR (https://www.ncr.com) estimates that, on average, this technology can reduce customer queue times by 40%. From the customer's point of view, if stores minimise labor costs, self-checkout can also reduce costs for consumers. Self-checkout can provide a higher-quality consumer experience if employees are reassigned to other tasks. According to the data of an American chain grocery store, after the implementation of self-checkout, 10% of their sales came from self-checkout. They were able to transfer 7% of the front-end labor force to other stores operations [18]. Tesco has also invested heavily in self-checkout technology in its retail stores.

2.2. Development and Feasibility of New Technologies. For self-service technologies, the need for appropriate technology to improve the efficiency of the shop and optimize the consumer shopping experience has become an important issue. In this study, this problem is divided into two parts: one is the association with the goods and the other is the capture of the perceptions of people and finally their systematic design to provide solutions to the above problems. RFID is thus an important technology for correlating goods, and the deep learning-based recommendation algorithms used in online shopping are important for capturing and outputting human perceptions.

RFID and related automatic identification systems are designed to solve the problem of electronic tag technology. This type of technology allows the use of electromagnetic challenge/response exchange. This kind of technology can automatically identify objects, places, or people at a distance without being able to remove them directly (Want, 2004). Kinsella (2003) described RFID as a simple technology. This technology enables machines to share information wirelessly. While the idea behind the technology is relatively simple, it is a system made up of tags attached to the products that send the information, card readers that receive the information the tags send, and software that collects and stores the information correctly. Both the tag and the reader are connected to an antenna that sends and receives data between the tag and the reader. Once the tag transmits its information to the reader, the reader's job is to send it to the appropriate computer device (Boyle, 2002). RFID is likely to be the next technology used on a large scale in the retail environment. RFID tags are designed to help reduce theft and better locate items. The use of labels can increase customer service, match supply to product demand, and speed delivery. Unlike bar codes, which must be passed before a scanner, RFID tags can be read remotely from a device 20 feet away. This flexibility opens up many new ways for retailers to increase CRM. RFID can also speed returns, manage warranties, and provide after-sales support. Users can also calculate costs based on RFID, especially how expensive tags are now. RFID technology can track sales after they are made.

On the other hand, in recent years, deep learning has started to be widely used in areas such as recommendation. Various deep models are able to learn discernible features from unstructured data through training. For example, Davidsan et al. proposed the "Item-KNN," which recommends items similar to items previously visited by the user, and the similarity between objects and items are expressed by the number of cooccurrences of the two [21]. Rendle et al. proposed FPMC for the following basket recommendations: FPMC uses matrix decomposition and first-order Markov chains to capture users' long-term preferences and short-term transitions, respectively [22]. BPR-MF is one of the most widely used matrix decomposition methods, which optimizes a pairwise ranking objective function by stochastic gradient descent to address the inability of traditional matrix decomposition methods to be directly applied to session-based recommendation tasks [23]. A mixed model of a shallow linear model and a deep feedforward network model has been trained, combining the memory capability of the shallow model and the generalization capability of the deep model into one, thereby balancing the accuracy and scalability of the recommendation system [24]. HRNN-Init is a personalized recommendation approach based on GRU4Rec and adds an additional GRU layer to model the evolution of user interests across sessions [25]. Zhou et al. propose a local activation unit for assigning different weights to users' historical behaviors so as to adaptively adjust the degree of influence of various historical behavioral features on the final results according to the candidate products [26, 27].

These neural networks (e.g., RNN, GRU, LSTM, and CNN) have significant advantages in modelling sequential data and have been widely used in personalized recommendations. Reference [28] pioneered the application of RNN to personalization, and their proposed GRU4Rec model showed a significant improvement over traditional methods. The CNN-based serialization recommendation embeds the items the user has interacted with into an $n^{*}k$-dimensional potential matrix and is treated as an image for processing. For example, the convolutional sequence embedding recommendation model proposed in the literature [29] consists of an embedding layer, a convolutional layer, and a fully connected layer. The convolutional layer consists of a horizontal convolutional layer and a vertical convolutional layer. In the horizontal convolution layer, all convolution results of the kF filter are maximally pooled to capture the most distinct features extracted by the filter by taking the most significant value. The operation of the vertical convolution layer is similar to that of the horizontal convolution layer. Finally, the outputs of the two convolutional layers are connected, and the tightly connected neural network layers are fed in to obtain higher-level features.

The attention mechanism affects the output by assigning different weights to the inputs. Several studies have attempted to use attention mechanisms to improve the performance and interpretability of serialized recommendations [16, 30, 31]. Reference [31] proposed a RIB model that introduces attention mechanisms into RNN-based serialized recommendation models. Reference [17] introduced an item-level attention mechanism in a local encoder model to capture the user's main intent in the current session, allowing different parts of the sequence to be dynamically and selectively input by the decoder. Besides, RNN-based recommendation models cannot model association relations at longer distances, but adding a self-attention mechanism inside the sequence can model across distances, and purely self-attention-based models (without any convolutional or recursive operations) started to be applied to the personalized recommendation. Reference [18] built a recommendation model based on a self-attention mechanism.

A review of RFID and a further review of the deep learning-based neural networks and attention mechanism reveals that both have been continually optimized in their respective fields and terms of accuracy and model, but that no research or offline consumer-oriented testing activities have been carried out to combine the technologies, while their principles and technologies show good potential for cross compatibility.

3. Design Rationale and Implementation

We designed a smart self-service kiosk that allows consumers to autonomously purchase goods in a store without a cashier or security. We leverage insights from operational patterns, and we obtain actual quantitative data from three convenient stores (e.g., 711 and FamilyMart) in China to better understand the degree and origin of the SST in a retail setting. Then, we translate what we learn in conceptual design (as illustrated in Figure 1) and implement a corresponding artefact that consists of a structure that is physically similar to a container, a WeChat-based shopping application, an in-store self-checkout accouter, and a backend operational management system. The functionalities involved are mainly grouped as follows: (1) express

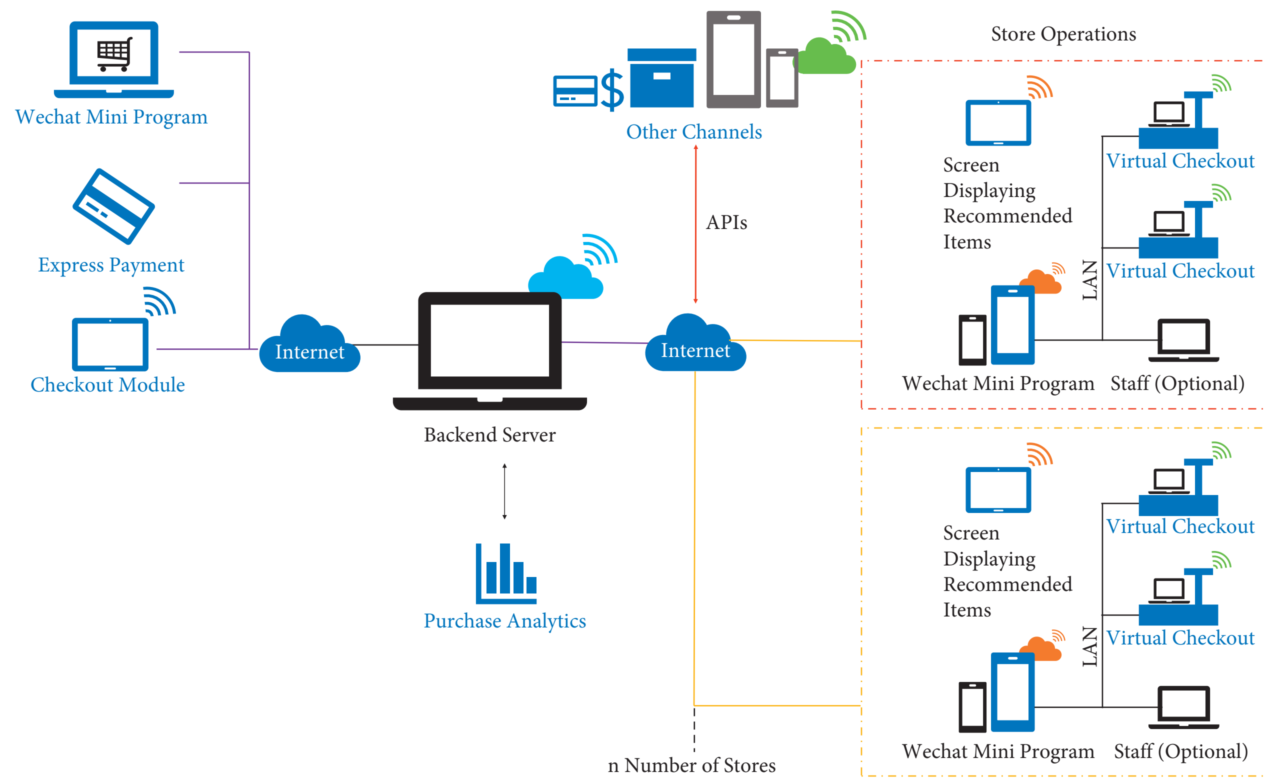

FIGURE 1: An overview of the system architecture of Kkbox system.

commodity checkout module, (2) user identification, and (3) operational support and shopping assistance.

3.1. Physical Design. The physical design of the kiosk requires local adjustments. Adjustments to the physical structure of the kiosk were made to accommodate the limited internal space and the SST setting. The correct placement of RFID tags is the key to successful data reading. As a result, the checkout counters and merchandise organization design will also be redesigned. The store's visual marketing identity and near transparency to the consumer should be an equal priority. Failure to follow this principle will reduce consumer comfort in an SST environment. In addition, we had to consider input/output controls for kiosks, self-checkout processes, and other parts of the organization that fit into the new setup. Based on the abovementioned considerations, a commercially available retail kiosk is designed and manufactured, as illustrated in Figure 2.

3.2. Express Commodity Checkout Module. Kkbox's checkout module design is shown in Figure 3. RFID tags are attached to goods. Each RFID tag contains information corresponding to a unique serial number associated with the operations support system. The server's database stores the name, unit price, manufacturer, and related data of all goods.

UHD RFID could be used to read all tag information simultaneously. It can scan up to 1,000 items at one time. Equipped with an upper RFID four-wire spiral antenna, the reading distance can be extended to 10 meters. RFID readers and antennas are placed near the terminal's main entrance in the checkout area. When the goods pass through the checkout area, the reader placed in the checkout area can read the information on the RFID tag attached to the goods in a short time and transfer this information to a computer, which will retrieve the goods identified by the serial number in the goods database. It then shows the customer a list of items and the total price for further payment confirmation. The checkout page will pop up after the relevant processing process is complete.

In such a way, all the commodities can be scanned in less than 10 seconds (even if there are many commodities, for example, 100 pieces) without taking them out of the shopping bag. Besides, there will exist no queues and no access issues but 24 hrs-opening hours. Kkbox enables flexibility in stock management, inventory control, and accurate good trend analysis.

The payment system is implemented based on Alipay and WeChat Pay as an essential component, because WeChat Pay is now the most popular payment method in China, linking features such as face recognition, and it is already being used on shopping mall SST. During a customer's stay in the kiosk, his WeChat Pay account will be used as their identifier. The critical function of the fast payment module is to purchase physical products from physical retailers without interacting with the POS machine. This immediate payment process consists of four main components: (1) a user presents their WeChat Pay at the kiosk entry for authentication. (2) Users walk into the

Figure 2: A Kkbox kiosk deployed in a community in Taizhou, China.

checkout area to check the total amount and present the code for payment. (3) Our operation system initiates the payment process to payment gateways of WeChat Pay. The due amount is charged to the payment account. (4) The payment results will be sent back to the user's account and stored on the user's mobile phone or smartphone (as illustrated in Figure 3).

3.3. Shopping Assistance and Item Recommendation. Many retailers have developed "shopping assist" technology to enhance the online and in-store shopping experience. This technology can help filter and narrow down the range of products available. In addition, the technology also allows for an in-depth comparison of selected products [16]. WeChat is the most popular application, and there are over 100 million users in China. Kkbox implements a WeChat-based application (as shown in Figure 4) for shopping assistance and self-checkout, facilitated with item recommendations. It applies to the setting of administrators, arrangement of duty schedules, management of customer and member information, and control of supermarket income information. Online shopping assistance has been shown to reduce search costs and increase convenience and the quality of purchase decisions [16, 29]. Consumers increasingly utilize mobile phones in the shopping process—primarily for information search in the prepurchase phase and less for actual purchase transactions [32–35].

To achieve the item recommendation, a recommendation model (as shown in Figure 5) is adopted to achieve the personalized recommendation of goods. The model combines the multilayer self-attentive network with the AUGRU [21–24, 36]. It extracts the user's interest vector from the whole user's sequence behavior through the multilayer self-attentive network structure and AUGRU and then matches the sequence to be recommended with the user's interest through the bilinear interpolation matching function to finally obtain the sequence of goods that the user may interact with.

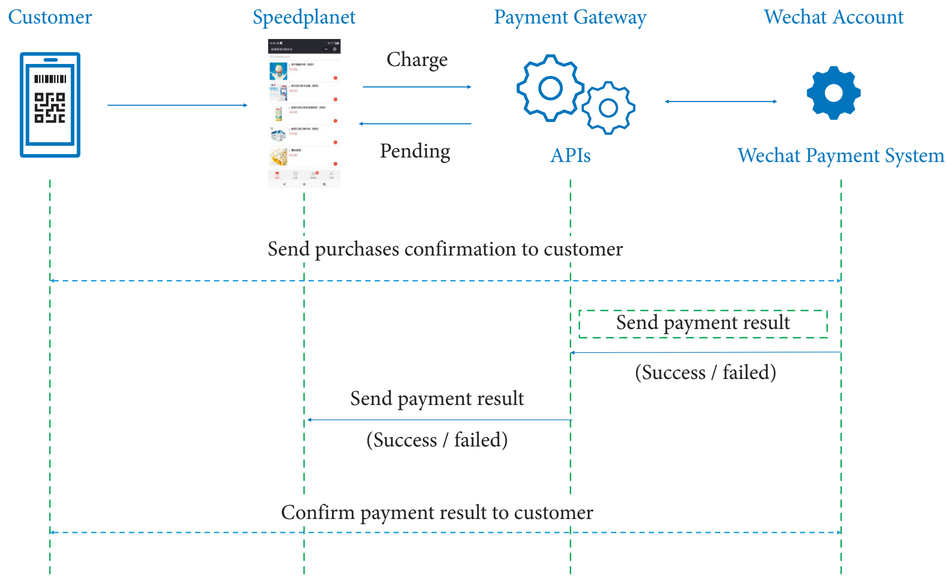

Figure 3: A fully automatic payment workflow.

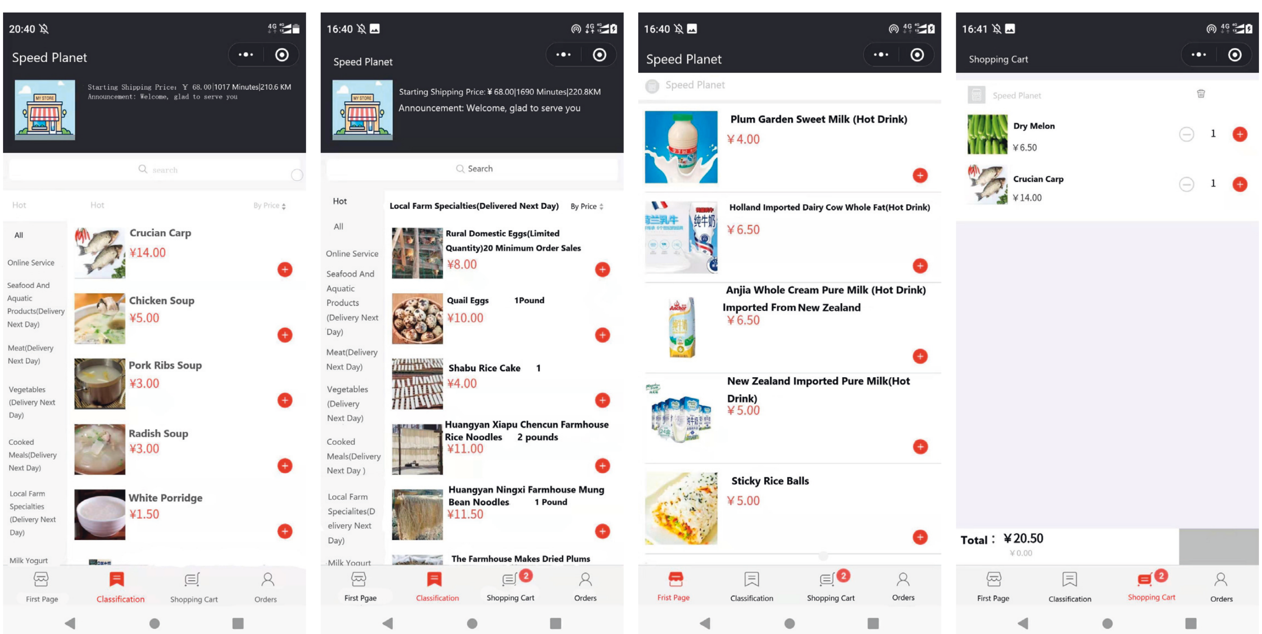

Figure 4: Screenshots of WeChat built-in app view of (1) available commodity list, (2) promotional commodity list, and (3) online shopping checkout and confirmation.

The model represents the sequence information of the goods by embedding the location codes of the historical goods, ensuring that the user's old points of interest neither disappear over time nor create unnecessary interference. The model uses a multilayer self-attentive network mechanism for the relationships between users' historical goods. The input to the encoder consists of two components: the user's historical purchase item characteristics and the purchase order of historical shopping items. By introducing location coding in the coding process, the sequence information in the user's history of clicking on goods can be effectively mined, thus enabling more effective modelling of the user's interests.

In this study, we will use a multilayer self-attention network to model the intrinsic relationship between users' historical interactions. Each layer of the self-attention network contains a query, key, and value. Their values are identical and are all vector representations of the user's historical shopping product feature vectors formed after the embedding layer. In a layer of self-attentive network, the query and key computation processes are described as follows:

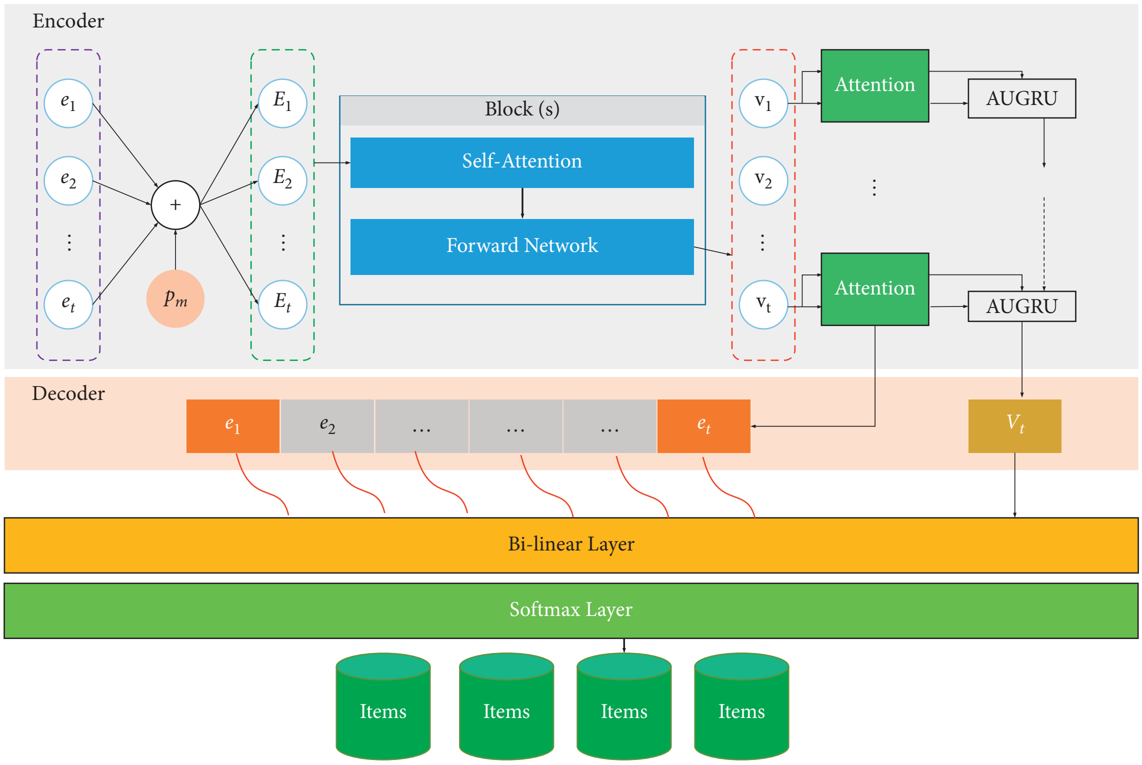

Figure 5: The insight of the item recommendation module.

$$\partial_t = \text{soft max}\left(\frac{QK^T}{\sqrt{d}}\right). \quad (1)$$

Finally, based on the obtained similarity matrix, a weighted average of each item in the historical interaction list with the vector of all historical interaction items for that user is obtained, and the representation vector of each item $\{v_1, v_2,..., v_t\}$ is as follows:

$$v_t = \sum_{i=1}^{t} \partial_i E_i. \quad (2)$$

To give the model nonlinearity and to consider the interaction between different dimensions, we input the representation vector of each clicked commodity into the feedforward network:

$$F_t = \text{ReLU}\left(v_t m + p\right) n + q, \quad (3)$$

where ReLU () is a nonlinear function, d is a scaling factor to adjust the inner product size, m and n are learnable linear matrices, and p and q are d-dimensional bias vectors. After one layer of self-attentive network, the sequence $\{v_1, v_2,..., v_t\}$ basically aggregates the embedding of the user's historical interaction goods; however, due to the diversity of goods attributes and the complexity of user interactions, it is necessary to use one or several more layers of the self-attentive network to capture more complex intrinsic relationships between interaction sequences, so we propose a multilayer self-attentive network for deeper relationship mining in the relationship extraction encoder. The kth block of self-attentive modules $(k > 1)$ can be defined as follows:

$$\begin{aligned} v_t^{(k)} &= F_t^{(k-1)}, \\ F_t^{(k)} &= \text{FFN}\left(v_t^{(k)}\right). \end{aligned} \quad (4)$$

With the multilayer self-attentive network, the model can not only capture the relationship between every two users' historical interaction goods but also focus "attention" on the goods' attributes that can really express users' interests by giving different attention scores to different goods' attributes, so that the model has the ability to capture more complex user interests. This allows the model to capture more complex user interests and facilitate the expression of user interests.

The attention function in this module is defined as follows:

$$\alpha_t = \frac{\exp\left(v_t m s_a\right)}{\sum_{j=1}^{T}\left(v_j m s_a\right)}, \quad (5)$$

where s_a denotes the embedding vector of the item to be recommended, the attention score reflects the correlation between the item to be recommended s_a and the input v_t, and a higher score indicates a stronger correlation between the two variants.

AUGRU adopts attention scores to update the state of the hidden layer of GRU to build the model of the dynamic interest of users. The bilinear interpolation matching function is used to match the user interest representation and the commodity to be recommended because it saves time and space compared to traditional methods. Given the goods that the user interacted with before moment t, the task of this model is to predict the goods that the user is about to interact with at moment $t + 1$. In the training process, we use the ADAM to update the model parameters, and the loss function is described as follows:

$$\mathrm{Loss}(g, u) = -\sum_{i=1}^{n} g_i \log(u_i), \quad (6)$$

where u is the true distribution of user-click sequences and g is the probability distribution of the model output results. Although the multilayer self-attentive network module can work towards the association relationship, as the number of network layers increases, the model is likely to be overfitted on the one hand. Moreover, as the number of layers increases, more parameters need to be trained, which will lead to more time and space required for the training of the model. To solve the above problem, we optimize the procedure as follows:

$$\begin{aligned} f(x) &= x + D(f(L(x))), \\ L(x) &= \gamma * * \frac{x - \lambda}{\sqrt{\mu + \varepsilon}} + \eta, \end{aligned} \quad (7)$$

where x denotes all feature embeddings for each item, $f(x)$ denotes a self-attention layer, $**$ denotes the product operation, λ represents the mean, μ denotes the variance of the feature embeddings, γ is the scaling factor, η is the bias term, and ε is a floating point number to prevent the situation that divisor is zero, respectively. That is, when using a multilayer self-attentive network, the input of each layer will be executed by a normalized layer, then the output of that layer is randomly deactivated, and the input x is added to the final output to prevent information omission.

4. Result

4.1. Experimental Setting. In this experiment, we collected the data from 6 groups of self-service kiosks for community retailing deployed in Taizhou, China. We selected a kiosk deployed in a high-density community for observation. The experimental data including total visit, customer, total amount, in-store purchase, online purchase, and shopping time are taken into consideration to evaluate the efficiency and acceptance of our proposed solution. The data collected range from July to August 2018.

Furthermore, we distribute opening flyers to passersby during the illustrated commuting hours through the part-time staff (e.g. students). We also deliver ads in WeChat moments to promote our service account on WeChat platform according to the user's location. If a customer purchase items through the app, the items would be sent to their home in one hour, but an extra delivery fee would be charged according to the purchase amount.

4.2. Conversion and Usage: Online Service Is More Attractive to Customer. Of the 500 users identified as eligible samples, 202 have purchased at least one product. There is no financial incentive for users to use the application. A total of 140 users made at least one transaction through the WeChat-based shopping app, while 26 users used the self-checkout app multiple times when purchasing products. As a result (illustrated in Figure 6), about 56% of buying users make more than one transaction. Conversion rates and utilisation rates are higher considering the recruitment channels of clients. We distinguish between proximity shopping (via flyers in stores) and WeChat online shopping. Queue types can be assigned to users based on their respective registration dates. All online customers have a higher relative conversion rate than in-store customers. Online promotions contribute more to business than regular in-store promotions.

Of all selected customers, 48 percent choose to purchase products through WeChat. Thirty-one percent prefer to use these two channels to buy products. About 38% of all face-to-face hires made purchases using the app, compared with 15% in the remote recruitment group. More impressively, 46% of face-to-face recruits cited "saving time" as their primary motivation for buying, and even 57% of those who used the app to purchase if they were in a convenience store "every day"—all (8) of those users made at least one more purchase through the app. The transactions vary in each kiosk due to their locations. However, we report an average total of 129 transactions issued in each kiosk per day. In our total sample, the most active user made a total of 31 transactions during the week, while the second most active user made a total of 26 transactions.

4.3. Time Issue. Aiming to analyzing the effects of increasing store rush and queues on the mobile app users, we analyze the time used to complete the purchase process from (1) time used to select products, (2) time for queuing if needed, and (3) checkout and payment (it is almost fixed) and its distribution during peak and nonpeak hours for all the transactions.

An in-store customer spends an average time of 107 seconds to complete the checkout and payment process (not including the time used to select commodities in the kiosk). Meanwhile, an app customer only spends an average time of 30 seconds to complete these steps but with a cost of extra charge for delivery. Comparatively, an average time of 4 minutes is required to complete the same process in a convenient store. Furthermore, we find that the mean purchase time from the app is about 390 seconds. For in-store mode, the average time used to select an item is 130 seconds. Meanwhile, the time for an app customer is over 400 seconds. We think there might be a process of making a comparison with another online e-commerce platform. Thus, the average time is longer than the in-store customer.

To compare our two metrics of the day, we relied on regular transaction data from these kiosks. We split trading into three different periods over 24 hours, with the morning peak from 1 a.m. to 7 a.m., the afternoon peak from 6 pm to 12 pm, and the rest of the day. Our results show that the mean and median of the two measures are almost equal or equal at different times of the day. We assume that the median shopping time in the afternoon and evening is slightly higher because there is less time pressure and users are more "strolling" shopping.

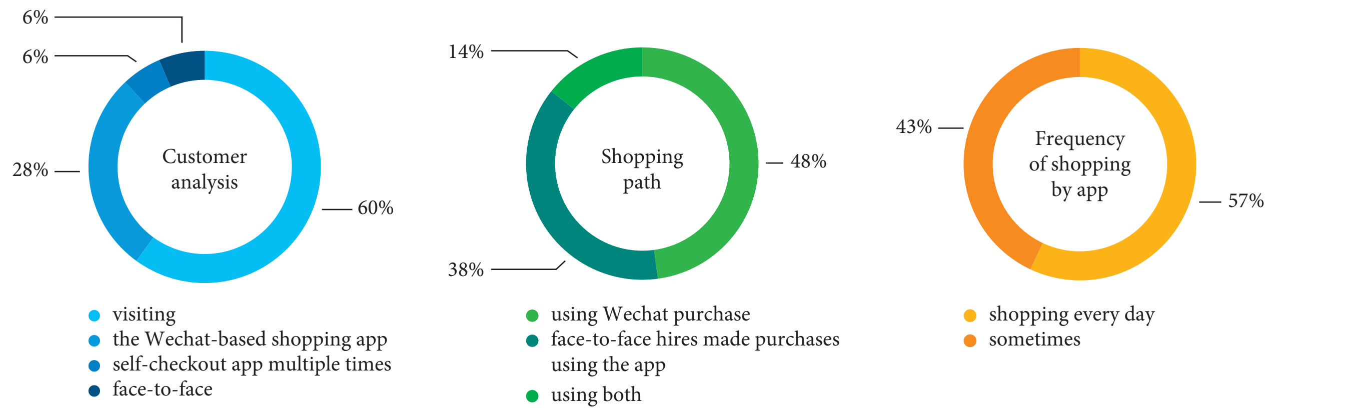

FIGURE 6: Analysis of customers' behaviors.

In order to avoid a broad calculation of time metrics that can be influenced by some product-specific factors, we define the calculation of the actual shopping time from the first shopping scan, which can be more accurate. We counted and calculated how long all in-store users spent shopping during the morning rush hour. The calculated average is used as a baseline and compared to the application users in the same morning session of the day. We made a total of 95 observations. The sample showed that in-store customers spent an average of 130 seconds buying, while app users spent an average of 30 seconds during the morning rush hour.

All these stores are characterized by lower unit prices and smaller items, including about 3 to 5 items per transaction and a purchase amount between 16 and 50 CNY. The share of alcohol and tobacco is 18% and 21%, respectively. The popular items include soft drinks, beer, tobacco, snacks, and bread and pastries. Our assessment also aims to understand consumer spending patterns in the current retail setting. For the pilot store, we selected that it has a central location and will be visited by more diligent customers; two of the three stores showed peak demand characteristics around working hours. The kiosk has the highest number of transactions in two periods (i.e., 7 a.m.–9 a.m. and 5 p.m.–7 p.m.). We observe that the increase during peak hours is usually more than three times the average demand. Kiosks, in particular, have long queues during peak hours, and customers face long waits. Therefore, we believe that our self-checkout module minimises the time and effort required by customers, even during peak hours, with long queues.

4.4. Item Recommendation Evaluation. We divide the data set into a training set and test set according to the time metric, in which 60% of the data is used as the training set, and the remaining 40% is further divided into 4 test sets randomly. For the evaluation of the recommendation system, Recall is an important metric, which counts the frequency of items in the test set that users actually clicked in the top K items in the recommendation list, and recall can be defined as follows:

$$\text{Recall}(K) = \frac{1}{F}\sum m_{\text{hit}}, \tag{8}$$

where K is set to 10, F is the total number of recommendations, and m_{hit} is the indicator function. If the target item appears in the current recommendation list, the value of the indicator function is 1, and vice versa is 0.

The experiments use a three-layer self-attentive network with an Adam optimization function. Table 1 lists that the model proposed in this study achieves better performance on all five datasets, indicating the effectiveness of the proposed approach. In addition, we can observe that the recommendation performance of the above models on Dataset2 is generally better than that on the other datasets, and a possible explanation for this is that Dataset3 has more users and interactions than the different datasets. At the same time, the number of items is relatively small.

As the number of self-attentive layers increases, the more thoroughly the model explores the relationships between sequences, the higher the recommendation performance. However, as the number of layers increases, the overfitting problem of the model to the training data and the excessive time and space consumed for training also occur. By observing the experimental results (in Table 2), we can learn that adding the self-attentive network to both datasets improves the evaluation indexes substantially than not adding the self-attentive network, which indicates that modelling the intrinsic relationship between user history interaction sequences is beneficial to improve on the same dataset, and the optimal performance of the model is achieved when the number of layers of the self-attentive network is 3 or 4, which means that the hierarchical self-attentive structure helps to learn more complex relationships. When the number of layers of the self-attentive network exceeds 4, the metrics of the model reach a plateau, which is due to the overfitting of the model.

Table 1: The comparison of recall results for different algorithms.

Experimental dataset	Wide and deep	DIN	FPMC	Item-KNN	BPR-MF	Proposed method
Dataset 1	0.7802	0.819	0.3901	0.2132	0.2309	0.7285
Dataset 2	0.6745	0.438	0.2904	0.4034	0.1035	0.8732
Dataset 3	0.7974	0.7252	0.1392	0.2972	0.2427	0.8029
Dataset 4	0.7189	0.5013	0.2158	0.4822	0.2521	0.8561

Table 2: The relationship between the number of the attentive network layer and algorithm performance.

Experiment group	Attentive network	Layer number	Performance
Group 1	Not engaged	0	0.1394
Group 2	Engaged	1	0.4224
Group 3	Engaged	2	0.5324
Group 4	Engaged	3	0.8013
Group 5	Engaged	4	0.8561
Group 6	Engaged	5	0.6332

5. Conclusion

We design and implement a self-service retail kiosk solution and evaluate its acceptance and practical use in a high-density community in Taizhou, China. Conclusions drawn from the usage records of 500 customers illustrate the positive value and consumer acceptance of the new retail kiosks. An entry survey provides more insight into the demographics and motivations of our study participants. It illustrates that almost half of users are often (at least once a week) unable to make purchases due to time pressures and long lines. For this, we compare it to the baseline time performance of the average in-store customer during peak hours, and the results show that the average customer saved 60 seconds by shopping at the kiosk. In addition, we can demonstrate that the purchase time required by application users was stable throughout the day, with delivery times prolonging even when queues occurred during peak morning and afternoon hours.

This study provides comprehensive guidance to better understand how to design a totally innominate self-help retail kiosk program and integrate mono in-store purchase with mobile retailing and instant delivery. This innovative practice model combines attention algorithms with RFID systems to project the technology of online product recommendations to customers into offline retail, meeting the need for more for fast consumption. Contemporary customers are basically popularised with smart terminals, which is an important basis for technological upgrades in SST, a sign of the era, and an important catalyst point in the study of smart communities that enables more ordinary residents to feel the convenience brought by the smarting of their lives.

The attempt on the algorithm is of great significance, but there are still limitations in the application research. It is suggested that future research can be carried out from the promotion of the product and the direction of multipoint multidata in line with the ability and arithmetic power of computing to optimize the research for user experience upgrading. In order to gain more understanding of application adoption and usage, as well as general consumption patterns in the current retail setting, our goal is to further expand our research with more participants over a more extended period of time and collect more data on mobile application user satisfaction.

Data Availability

The data are available from the corresponding author upon request (20310231@tongji.edu.cn).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Ministry of Housing and Urban-Rural Development (2021-K-148), General Topics of the Shanghai Philosophy and Social Science Programme (2021BCK004), and Soft Science Research Project of Shanghai 2022 Science and Technology Innovation Action Plan (22692106800).

References

[1] P. Singh, M. Gupta, A. Kumar, P. Sikdar, and N. Sinha, "E-grocery retailing mobile application: discerning determinants of repatronage intentions in an emerging economy," *International Journal of Human-Computer Interaction*, vol. 37, no. 19, pp. 1783–1798, 2021.

[2] K. Kallweit, P. Spreer, and W. Toporowski, "Why do customers use self-service information technologies in retail? The mediating effect of perceived service quality," *Journal of Retailing and Consumer Services*, vol. 21, no. 3, pp. 268–276, 2014.

[3] F. Bielen and N. Demoulin, "Waiting time influence on the satisfaction-loyalty relationship in services," *Managing Service Quality: International Journal*, vol. 17, no. 2, pp. 174–193, 2007.

[4] K. van Ittersum, B Wansink, J. M. E. Pennings, and D. Sheehan, "Smart shopping carts: how real-time feedback influences spending," *Journal of Marketing*, vol. 77, no. 6, pp. 21–36, 2013.

[5] X. Zhou, Y. Li, and W. Liang, "CNN-RNN based intelligent recommendation for online medical pre-diagnosis support," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 3, pp. 912–921, 2021.

[6] H. T. Cheng, L Koc, J. Harmsen et al., "Wide & deep learning for recommender systems," in *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems*, pp. 7–10, Boston MA, USA, September 2016.

[7] M. Quadrana, A. Karatzoglou A, B. Hidasi B, and P. Cremonesi, "Personalizing session-based

recommendations with hierarchical recurrent neural networks," in *Proceedings of the Eleventh ACM Conference on Recommender Systems*, pp. 130–137, Como Italy, Auguest 2017.

[8] Z. Cai and X. Zheng, "A private and efficient mechanism for data uploading in smart cyber-physical systems," *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 2, pp. 766–775, 2020.

[9] G. Zhou, X. Zhu, C. Song et al., "Deep interest network for click-through rate prediction Proceedings of the 24th," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, no. 7, pp. 1059–1068, London, UK, July 2018.

[10] X. Zhou, W. Liang, W. Li, K. Yan, S. Shimizu, and K. I. K. Wang, "Hierarchical adversarial attacks against graph neural network based IoT network intrusion detection system," *IEEE Internet of Things Journal*, p. 1, 2021.

[11] M. Shahid Iqbal, M. Ul Hassan, and U. Habibah, "Impact of self-service technology (SST) service quality on customer loyalty and behavioral intention: the mediating role of customer satisfaction," *Cogent Business & Management*, vol. 5, no. 1, p. 1, 2018.

[12] M. Hyoungeun, L. Heejung, and H. Heesup, "Self-check-in kiosk quality and airline non-contact service maximization: how to win air traveller satisfaction and loyalty in the post-pandemic world?" *Journal of Travel & Tourism Marketing*, vol. 38, no. 4, pp. 383–398, 2021.

[13] S. Gupta, S. Modgil, A. Gunasekaran, and S. Bag, "Dynamic capabilities and institutional theories for Industry 4.0 and digital supply chain," *Supply Chain Forum: International Journal*, vol. 21, no. 3, pp. 139–157, 2020.

[14] X. Zhou, W. Liang, K. I. K. Wang, and L. T. Yang, "Deep correlation mining based on hierarchical hybrid networks for heterogeneous big data recommendations," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 1, pp. 171–178, 2021.

[15] Z. Cai, Z. Xiong, H. Xu, P. Wang, W. Li, and Y. Pan, "Generative adversarial networks," *ACM Computing Surveys*, vol. 54, no. 6, pp. 1–38, 2021.

[16] W. Liang, Y. Hu, X. Zhou, Y. Pan, and K. I-Kai Wang, "Variational few-shot learning for microservice-oriented intrusion detection in distributed industrial IoT," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 8, pp. 5087–5095, 2022.

[17] A. Davies, L. Dolega, and D. Arribas-Bel, "Buy online collect in-store: exploring grocery click&collect using a national case study," *International Journal of Retail & Distribution Management*, vol. 47, no. 3, pp. 278–291, 2019.

[18] M. Faraoni, R. Rialti, L. Zollo, and A. C. Pellicelli, "Exploring e-Loyalty Antecedents in B2C e-Commerce," *British Food Journal*, vol. 121, no. 2, pp. 574–589, 2019.

[19] Z. Cai, Z. He, X. Guan, and Y. Li, "Collective data-sanitization for preventing sensitive information inference attacks in social networks," *IEEE Transactions on Dependable and Secure Computing*, vol. 15, no. 4, pp. 577–590, 2018.

[20] D. Li, C. Chen, Q. Lv et al., "An algorithm for efficient privacy-preserving item-based collaborative filtering," *Future Generation Computer Systems*, vol. 55, no. C, pp. 311–320, 2016.

[21] X. Zheng and Z. Cai, "Privacy-preserved data sharing towards multiple parties in industrial IoTs," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 5, pp. 968–979, 2020.

[22] L. Guo, J. Chen, S. Li, Y. Li, and J. Lu, "A blockchain and IoT based lightweight framework for enabling information transparency in supply chain finance," *Digital Communication and Network*, 2022.

[23] F. Lu, Z. Zhang, and L. Guo, "A lightweight hand-crafted feature enhanced CNN for ceramic tile surface defect detection," *International Journal of Intelligent Systems*, 2022.

[24] D. Castro, R. D. Atkinson, and S. J. Ezell, "Embracing the self-service economy," *SSRN Electronic Journal*, 2010.

[25] X. Zhou, X. Xu, W. Liang, Z. Zeng, and Z. Yan, "Deep-learning-enhanced multitarget detection for end-edge-cloud surveillance in smart IoT," *IEEE Internet of Things Journal*, vol. 8, no. 16, pp. 12588–12596, Aug. 2021.

[26] G. Zhou, N. Mou, Y. Fan et al., "Deep interest evolution network for click-through rate prediction," in *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 5941–5948, Honolulu Hawaii, USA, January 2019.

[27] Z. Cai, X. Zheng, J. Wang, and Z. He, "Private data trading towards range counting queries in internet of things," *IEEE Transactions on Mobile Computing*, p. 1, 2022.

[28] F. Liébana-Cabanillas, J. Sánchez-Fernández, and F. Muñoz-Leiva, "Antecedents of the adoption of the new mobile payment systems: the moderating effect of age," *Computers in Human Behavior*, vol. 35, pp. 464–478, 2014.

[29] G. Salton, A. Wong, and C. S. Yang, "A vector space model for automatic indexing," *Communications of the ACM*, vol. 18, no. 11, pp. 613–620, 1975.

[30] B. Adipat, D. Zhang, and L. Zhou, "The effects of tree-view based presentation adaptation on mobile web browsing," *MIS Quarterly*, vol. 35, no. 1, pp. 99–121, 2011.

[31] A. A. Alalwan, "Mobile food ordering apps: an empirical study of the factors affecting customer e-satisfaction and continued intention to reuse," *International Journal of Information Management*, vol. 50, pp. 28–44, 2020.

[32] J. Davidson, B. Liebald, J. Liu et al., "The YouTube video recommendation system," in *Proceedings of the fourth ACM conference on Recommender systems*, pp. 293–296, Barcelona, Spain, September 2010.

[33] X. Zhou, X. Yang, J. Ma, and K. I. K. Wang, "Energy efficient smart routing based on link correlation mining for wireless edge computing in IoT," *IEEE Internet of Things Journal*, May 2021.

[34] S. Rendle, C. Freudenthaler, and L. Schmidt-Thieme, "Factorizing personalized Markova chains for next-basket recommendation," in *Proceedings of the 19th international conference on World wide web*, no. 4, pp. 811–820, Raleigh North Carolina USA, April 2010.

[35] B. Loni, R. Pagano, M. Larson, and A. Hanjalic, "Bayesian personalized ranking with multi-channel user feedback," in *Proceedings of the 10th ACM Conference on Recommender Systems*, p. 361, Boston Massachusetts, USA, September 2016.

[36] G. Linden, B. Smith, and J. York, "Amazon.com recommendations: item-to-item collaborative filtering," *IEEE Internet Computing*, vol. 7, no. 1, pp. 76–80, 2003.