
Intelligent

Inventory Control System

using data mining Apriori algorithm

YuanWang A00258427

Thesis presented for the degree of

MSc in Software Engineering to the

Department of Software Engineering,

Athlone Institute of Technology.

Supervisor:

MAMOONA ASGHAR

April 30th, 2019, Althone, Ireland

Copyright Yuan Wang, 2019

Declaration

I hereby certify that the material, which is submitted in this thesis towards the award of MSc. Software Design, is entirely my own work and has not been submitted for any academic assessment other than part fulfilment of the above named award.

Future students may use the material contained in this thesis provided that the source is acknowledged in full.

Signed.....

Date.....

Abstract

With the continuous development of the digitization process in the world, inventory management has become more challenging as a business issue. It itself plays a role in promoting the rapid development of business. Convenience stores have always played an important role as an integral part of modern people's daily lives. And inventory management is the first big problem that needs convenience stores to face and solve. The main purpose of this project is to build digital inventory management system (inbound management, outbound management, in-store merchandise management, etc.) by using data mining algorithm such as Apriori algorithm for an intelligent inventory management system. The convenience store records the goods on a daily basis, and obtains the degree of correlation between the products, thereby rationally and effectively arranging the shelf positions of the products and improving the efficiency of inventory management systems along with customer satisfaction.

Acknowledgements

Firstly, I appreciate my creator for thus far he has led and guided me.

I wish to say a big thanks to all those that supported me throughout my MSc in Software Engineering ranging from my family, lectures, and friends.

I would like to take this opportunity to thank and appreciate my thesis supervisor Dr. Mamoon Asghar for all her assistance, guidance and patience in the design and development of the system over the course of the academic year. My supervisor, you have been wonderful. It was a great privilege having you to supervise me, you supported me, encouraged me and was always available.

Many Thanks to you all I really appreciate it.

Table of content

Chapter I	1
Introduction	1
1.1 Overview of Retail industry	1
1.2 Overview of Data mining and its application	2
1.3 Object of the thesis	3
1.4 Contribution of the thesis.....	4
1.5 Thesis Outline	5
Chapter II.....	6
Literate Review.....	6
2.1 Apriori Algorithm	6
2.1.1 Logic of Apriori algorithm.....	8
2.1.2 Apriori algorithm produces frequent sets	10
2.2 Inventory management System	14
2.2.1 Inventory management objective	15
2.2.2 Development of warehouse management system	16
2.3 Shopping behavior analysis	17
Chapter III.....	20
System Design	20
3.1 Overview of tools and technology.....	20
3.1.1 Front end.....	20
3.1.2 Back end	23
3.1.3 Database side	26
3.2 System architecture	27
3.2.1 Architecture	27
3.2.2 Execution logic(data flow chart)	28
3.2.3 Database E-R diagram	29
3.2.4 UML sequence diagram	30

3.2.5 Use case diagram.....	31
3.3 Requirement.....	31
Character IV	44
System implementation and Testing	44
4.1 Implementation.....	44
4.2 Testing	53
4.2.1 Unit Testing	53
Chapter V	59
Conclusion.....	59
5.1 Reflection.....	59
5.2 Recommendation and future work	60
Reference:	61

Chapter I

Introduction

The main target customers of this project are convenience stores and small supermarkets, all of which are affiliated with the retail industry. The retail industry plays an important role in our daily lives, and it plays an irreplaceable role in integrating production and consumption. At the same time, the development of the retail industry is also constrained by many other factors, such as supply chain management, the impact of the rapid development of e-commerce and so on. At the same time, economic uncertainty has a negligible impact on the retail industry. For the modern economic development, we are gradually transitioning from the seller's market to the buyer's market, and more and more companies are being eliminated by consumers. For the retail industry, in this context, how to provide customers with more reasonable and more humane services becomes more and more important. Promoting employees in their own business is one aspect. The other is to provide higher quality services by improving their management. Intelligent management of inventory is one part of it.

1.1 Overview of Retail industry

The retail industry occupies a very important position in the Irish economy. According to data from March 2017, ¹ Ireland's retail industry generated 7billion of tax revenue, accounting for 12% of total GDP, and the retail sector ranked fourth in Europe, according to the information mentioned in Retail Ireland MonitorQ3. Convenience stores and supermarkets have seen corresponding growth in all aspects in the third quarter of 2018. Compared with 2017 and 2018, the overall growth rate was 4.4%, compared with the same period in 2017, an increase of 4.0% in the third quarter of 2018. In September, September increased by 1.9% compared with August, and September 2018 increased by 4.4% compared with the same period in September 2017. It should be noted that this is still in the third quarter of the Christmas consumption season. For the Irish retail industry, its composition is mainly composed of small family businesses, active retail and wholesale companies, about 37,400, almost all over the country.

¹ The main data cited in this section comes from retail Ireland-- www.retailireland.ie

1.2 Overview of Data mining and its application

As the pace of digitization in the modern world is getting faster and faster, our world is gradually digitizing, and the things in our lives are being digitized. These data are collected in different fields to provide people with more convenient service. As a result, many emerging fields have emerged, and some traditional areas that rely on data have renewed their vitality. For example, economics, meteorology, management, scientific research and other traditional fields, through data to better analyze the problems in our existing society, expand our unknown areas, emerging areas include: big data analysis, electronic communications, E-commerce, etc., these areas rely on the information provided by the data to better provide services, and even for the government departments, through the analysis of the data, to better discover the problems that the country currently has in various fields, Then decide the next strategy. The importance of data is unpredictable in today's world, and many researchers are attracted to it and are engaged in the wave of large research data.

Data mining, also known as knowledge discovery (KDD), mainly refers to:

The science of extracting useful information from large data sets or database [1].

Data mining involves the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data sets [2].

However data mining itself is not a new technology: people have been analyzing data on computers since the invention of the first computer, and in the past few centuries, people have been analyzing data without a computer. Nowadays, with the digitization process in the world, data mining has become a rapidly developing discipline in the field of science.

Data mining generally refers to the process of searching for information hidden in it from a large amount of data through an algorithm. Data mining is often associated with computer science and achieves these goals through statistics, online analytical processing, intelligence retrieval, machine learning, expert systems (reliant on past rules of thumb), and pattern recognition.

Data mining and data analysis are difficult to distinguish. The biggest difference between data mining and conventional data analysis methods is twofold: First, the former has a higher degree of automation; second, the former has high applicability.

Problems that apply to data mining may have the following characteristics:

- Suitable for automated analysis using computers
- Involving a large number of related variables, each variable has a huge amount of data
- System characteristics are completely unknown or poorly understood
- It is difficult to carry out routine theoretical analysis through steps such as reasoning, calculus, hypothesis and simplification.

For the problems we are going to analyze, data mining is very suitable for this scenario. In the inventory management, the use of data mining techniques and algorithms can maximize the rationality of inventory management and more efficient distribution.

1.3 Object of the thesis

Since the 21st century, the pressure on warehouse management has increased, especially in the storage management of inventory. Since inventory control involves many aspects such as warehouse management, outbound management, inventory management, and commodity inventory, a large amount of business data is generated in these links, which contains a large amount of unknown and valuable information. In this regard, the inventory management of supermarkets is more prominent.

For convenience stores and small supermarkets, how to improve their management reasonably, a key point is to improve their inventory management, how to achieve more effective shelf arrangement for goods in the warehouse, so that consumers can purchase goods. When you get the goods you need faster, you can get a better shopping experience, which is undoubtedly the problem that this project wants to solve. The advantage of this is not only that the consumer gets a better shopping experience, but also enhances the management efficiency of the enterprise itself, and the convenient shopping experience brought by the new socialized customer relationship management² also brings a brand new Business opportunities, through the promotion of consumers who have experienced good service, will also bring more new customers and new

² Socialized customer relationship management means that enterprises can be integrated into social applications, such as Facebook's SCRM system, by centralizing brand promotion channels, customer management, and media delivery. Maximize the customer base that the brand may acquire, and achieve precise marketing and refined management with Facebook's powerful social attributes.

resources to the company itself.

Convenience stores and small supermarkets, as a daily retail industry, will generate a lot of business data every day, and the value behind these without data is not easy to estimate. Through data mining, we finally The goal is to have these data ultimately fed back to these convenience stores and small supermarkets to help them more rationally arrange inventory and manage the corresponding goods.

1.4 Contribution of the thesis

There are many kinds of warehouse system management software in the market, but most of the software often lacks in-depth analysis of data, just simple data management. As an intelligent inventory management software based on data mining algorithm, this project is aimed at the vast small retail enterprises. It deeply studies the value hidden behind the data and the inventory management model based on the target users. The main contributions are as follows:

- Product inbound management - According to the size of the small retail store, the product into the warehouse, and provide a reasonable warehousing query mode.
- Product outbound management - According to the daily data flow of small retail stores, to provide merchandise sales, merchandise outbound inquiries, and product sales ranking.
- Data analysis model based on Apriori algorithm - This program adopts the combination of Apriori algorithm and inventory management, sales mode, based on the sales volume of goods in inventory, and generates the relevance of goods to reasonably arrange the shelf position of goods, thereby improving inventory management efficiency and Product sales.
- Inventory Management - Provides the corresponding product inquiry mode. At the same time, according to the daily small retail store data traffic, it dynamically provides corresponding information prompts, such as insufficient stock products and products are about to expire so that users can formulate more reasonable sales strategies.

1.5 Thesis Outline

The organization of thesis in the form of chapters is given below:

Chapter 1: Project Introduction – shows information relating to the background problem of the retail industry and data mining.

Chapter 2: Literature Review – Shows an investigation into other papers and algorithm can use to improve the realization of the project. Includes research on Apriori algorithm, analysis of inventory management systems, and analysis of human shopping behavior.

Chapter 3: System Design – Shows the design of the application with information relating to the system architecture and requirements of the system including system architecture, data flow diagram, E-R diagram, UML sequence diagram, use-case diagram, and descriptions.

Chapter 4: System Implementation and Testing – This is the system implementation and testing of the application.

Chapter 5: Conclusion – It concludes the thesis and has some suggestions for future work.

Chapter II

Literate Review

Inventory Management (IM), as a business that cannot be ignored by any retail company today, plays an irreplaceable role in improving business performance and is the first priority for all retailers. Reasonable and efficient inventory management not only brings about the reduction of costs, but also the efficiency of warehouse use. In general, managers face a large inventory needs analysis and considerations, such as the cost of inventory, sales methods, supply models and so on. Faced with the data generated by a large amount of inventory, it is necessary to use the corresponding algorithm for analysis.

This chapter mainly focuses on discuss and review several different papers on how to implement intelligent inventory management and optimization issues. It will examine the following technologies:

- Apriori Algorithm
- Inventory management system
- Shopping behavior analysis

For each of the techniques mentioned above, the relevant papers I have cited are relevant arguments for their technical surprises to demonstrate the usefulness of the technology and the possibility of combining it with the techniques used in this article. This will verify the possibility of formal implementation of the project and the possibility of using verification in the real world after the project is completed.

Although it may not be as accurate as the papers mentioned above, the system being developed as part of this thesis will be cheaper to implement and be more useful in real world situation.

2.1 Apriori Algorithm

The Apriori algorithm is one of the world's top ten famous data mining algorithms. It is a classic algorithm for implementing frequent item mining. It uses the associated rules

to continuously expand the subset of the frequent items to obtain all the frequent items. Explain the association rules, which are the relationships between items and items found in large amounts of data [3]. The purpose of correlation analysis is to find interesting relationships from large data sets. These relationships can take two forms: frequent item sets or association rules.

- Frequent item sets are collections of items that often appear together.
- Association rules suggest that there may be a strong relationship between the two items.

In data mining, there are scanning methods of depth-first-search (DFS) [4] and breadth-first-search (BFS) [5]. The depth-first search means that the algorithm traverses the data set from top to bottom from left to right, and the breadth-first search is the algorithm traverses the data set from top to bottom from left to right.

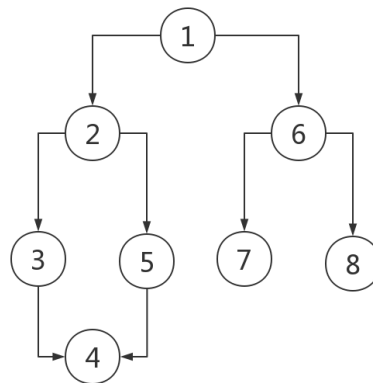


Figure2-1: depth-first-search chart

After accessing a starting vertex v , DFS starts from v and traverses all vertices from top to bottom, from left to right, as shown in Figure 1. The order of execution is: 1->2->3->4->5->6->7->8

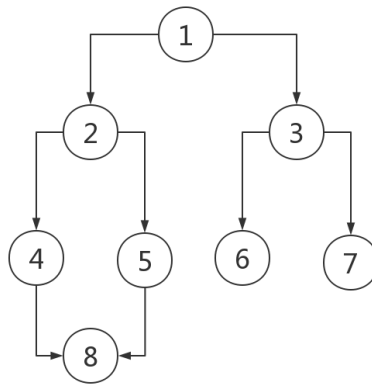


Figure2-2: Breadth-first-search chart

BFS is a node that traverses the tree (graph) along the width of the tree (graph) starting from the root node. If all nodes are accessed, the algorithm is aborted. As shown in Figure 2, The order of execution is: 1->2->3->4->5->6->7->8.

2.1.1 Logic of Apriori algorithm

The Apriori algorithm uses a breadth-first search algorithm to search and uses a tree structure to count candidate item sets. It first retrieves all the frequent item sets in the transaction database, and then uses the frequent item sets to construct rules that satisfy the minimum confidence³ of the users.

³ confidence($X \rightarrow Y$) = $P(Y | X)$. The minimum confidence, that is, the confidence greater than the condition, can be called a strong association rule. Usually, this value is dynamically set by the person according to different situations.

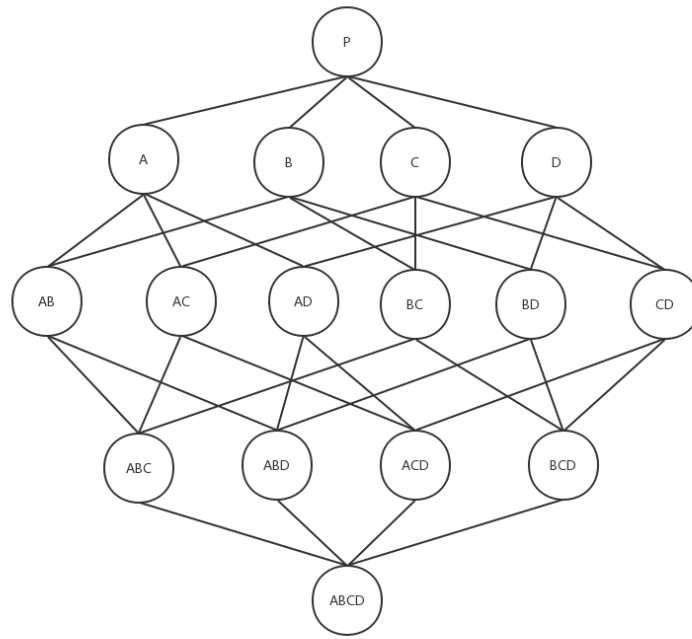


Figure2-3: Apriori algorithm execution logic diagram

From Figure 3, we can always see the execution logic of the a priori algorithm. Suppose there are 4 test items A, B, C, D. From P, the Apriori algorithm will find all the frequent sets in turn, and then remove the frequent item sets which are not satisfied with preconditions. This prerequisite is called minimum support⁴.

In this example, we found that the algorithm needs to iterate through the entire set to calculate the support for all combinations.

$$2^N - 1 = 2^4 - 1 = 15$$

As the number of items increases, the number of calculations grows exponentially. In the a priori algorithm, there is a characteristic that if an item sets is frequent, then all its subsets are also frequent. If an item set is an infrequent item set, then all its super-sets⁵ are also infrequent item sets.

⁴ support($X \rightarrow Y$) = $P(X \cup Y)$. The minimum support, that is, the number of occurrences of the same item set, may vary according to different situations.

⁵ If each element in a set S2 is in set S1, and set S1 may contain elements that are not in S2, then set S1 is a super-set of S2, which in turn is a subset of S1.

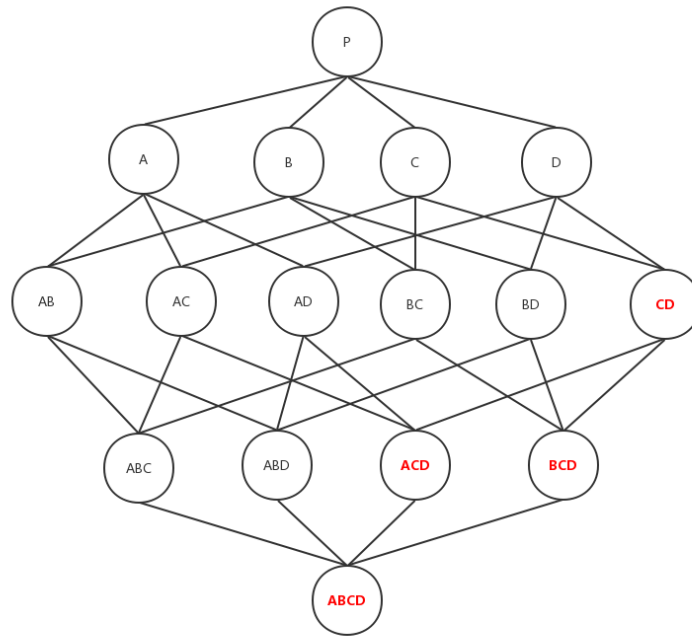


Figure2-4: Apriori algorithm execution logic diagram

In Figure 4, assuming that the red set $\{C, D\}$ is an infrequent item set, then we can know that $\{A, C, D\}$ $\{B, C, D\}$ $\{A, B, C, D\}$ are all infrequent of. In other words, the support of $\{C, D\}$ is calculated. If it does not satisfy the minimum support we set, we do not need to calculate $\{A, C, D\}$ $\{B, C, D\}$ $\{A, B, C, D\}$ support, because we know that these collections will not meet our requirements and thus avoid exponential growth in the number of item sets, thus calculating frequent item sets in a reasonable amount of time.

2.1.2 Apriori algorithm produces frequent sets

The Apriori algorithm is a classic association rule mining algorithm. To confirm whether the algorithm can be applied to the system, figure 2-5 shows the pseudo-code implementation of the algorithm [6].


```

Apriori( $T, \epsilon$ )
   $L_1 \leftarrow \{\text{large 1 - itemsets}\}$ 
   $k \leftarrow 2$ 
  while  $L_{k-1} \neq \emptyset$ 
     $C_k \leftarrow \{a \cup \{b\} \mid a \in L_{k-1} \wedge b \in \bigcup L_{k-1} \wedge b \notin a\}$ 
    for transactions  $t \in T$ 
       $C_t \leftarrow \{c \mid c \in C_k \wedge c \subseteq t\}$ 
      for candidates  $c \in C_t$ 
         $count[c] \leftarrow count[c] + 1$ 
       $L_k \leftarrow \{c \mid c \in C_k \wedge count[c] \geq \epsilon\}$ 
       $k \leftarrow k + 1$ 
  return  $\bigcup_k L_k$ 

```

Figure 2-5: Prior algorithm pseudo code implementation

In figure 2-5, where “ T ” represents the transaction sets, “ ϵ ” represents the minimum support, L_i corresponds to the frequent item sets obtained by the i th iteration, and C_i represents the candidate set for the i th iteration. The first line of pseudo code traverses all items in the transaction, counts the frequency of occurrences of each item, and retains all items that occur more frequently than the minimum support.

Based on the pseudo-code implementation, we can derive the basic generation steps of the frequent item sets of the Apriori algorithm [7].

Table 2-1: a database transaction record

TID	Items
T1	ABC
T2	ACD
T3	BCD
T4	ABCDE
T5	ACE
T6	BDE
T7	CD

The following steps are used to explain the generation of candidate sets and frequent item sets. *Itemset* represents the item set and *Sup.count* represents the support count. Table 2-1 transaction the minimum support is 4.

1) Scan transaction sets D, count the support count for each candidate, as shown in Table 2-2.

Table 2-2: Support count for candidate 1-item set options

<i>Itemset</i>	<i>Sup.count</i>
A	4
B	4
C	6
D	5
E	3

2) Compare candidate support counts with minimum support counts, delete candidates that do not meet the criteria, and generate frequent 1-item sets L1, as shown in Table 2-3.

Table 2-3: Frequent 1-item set L_1

<i>Itemset</i>	<i>Sup.count</i>
A	4
B	4
C	6
D	5

3) Generate candidate 2-item set C_2 from L_1 , and count each candidate, as shown in Table 4.

Table 2-4: Candidate 2-item set C_2

<i>Itemset</i>	<i>Sup.count</i>
AB	2
AC	4
AD	2
AE	2
BC	3
BD	3
BE	2
CD	4
CE	2
DE	2

4) Compare candidate support counts with minimum support and generate frequent The 2-item set L_2 is shown in Table 2-5.

Table 2-5: Frequent 2-item set L_2

<i>Itemset</i>	<i>Sup.count</i>
AC	4
CD	4

5) Generate candidate 3-item set C_3 from L_2 , and count each candidate, as shown in Table 2-6.

Table 2-6: Candidate 3-item set C_3

<i>Itemset</i>	<i>Sup.count</i>
ACD	2

In this example, since the minimum support is 4 when generating the frequent sets L_1 and L_2 , all the item sets smaller than 4 are deleted. This is called pruning that removes the frequent sets which is not satisfied with the minimum support. Finally, when the candidate set C_3 is generated since there is no frequent set satisfying the minimum support, the a priori algorithm stops here, and the frequent set L_3 is no longer generated. And then the calculation of the minimum confidence is started.

From the frequent set L_2 , we can conclude that there are two item sets that satisfy the condition, AC and CD, and then based on the formula for calculating the minimum confidence:

$$\text{Confidence } (X \rightarrow Y) = P(Y|X) = \text{Support_Count}(X \cup Y) / \text{Support_Count}(X)$$

We can conclude that:

$$\begin{aligned} A \rightarrow D: \text{Confidence } (A \rightarrow D) &= P(D|A) = \text{Support_Count}(A \cup D) / \text{Support_Count}(A) \\ &= 2 / 4 \end{aligned}$$

$$\begin{aligned} C \rightarrow D: \text{Confidence } (C \rightarrow D) &= P(D|C) = \text{Support_Count}(C \cup D) / \text{Support_Count}(C) \\ &= 4 / 6 \end{aligned}$$

When the confidence is greater than the minimum confidence, the rules can be called Strong association rules. In this example, if $\text{min_conf} = 60\%$, then $C \rightarrow D$ is a strong association rule.

2.2 Inventory management System

The warehouse management system is a product that emerges along with the digitization process. It replaces the original inefficient warehouse management, making the entire warehouse management more transparent, efficient, and controllable, improving the overall warehouse service efficiency.

2.2.1 Inventory management objective

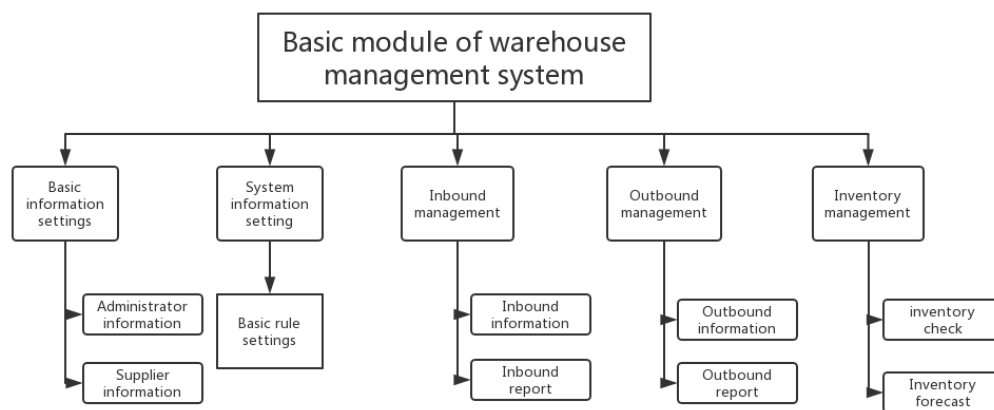


Figure 2-6: Basic Module of Warehouse management system

For warehouse management, there are many kinds of existing warehouse management systems, but the basic functional modules of the WMS (warehouse management system) are similar and the purpose is to improve warehouse management. We can divide the main purposes of warehouse management into the following point [8]:

1) Establish basic material information of the goods, including the name of the goods, the code of the goods, the bar-code or QR code, the picture information of the goods, the specification and model, the reference price of the goods, the unit of measurement, the unit of conversion, the initial inventory data, the initial warehouse, the unit price, Basic information such as inventory amount.

2) Out of the warehouse management, after the goods arrive, the warehouse management system scans the QR code or bar-code of the goods to scan the code into the warehouse, scan the code to check the quantity of the storage, the system records the inbound amount, time, and the delivery person, the handler, warehousing date, delivery unit and other information. Outbound is the same to establish a shipper, scan the goods bar-code or QR code according to the shipper, confirm the quantity of the

goods out of the warehouse, record the outbound customers, and the time of delivery.

3) The inventory is regularly counted in order to ensure the accuracy of the inventory quantity. Warehouse management personnel should regularly use the cloud warehouse for data inventory, update the inventory quantity in real time, correct incorrect book data, and ensure the accuracy of data in warehouse management systems and the real-time nature of data.

4) The data transfer, if the warehouse is out of stock, you need to call a certain part of the inventory from the B warehouse, should now submit a transfer application in the warehouse, the audit authority to review the equivalent transfer application, after the review. A warehouse manager can transfer the goods to the management staff of the B warehouse.

5) data analysis and summary, at the end of the month, the need for various financial statements, summary tables, schedules, customer statements, and other data analysis statistics, you can directly use the cloud warehouse to directly analyze the statistics through the input of various analysis statistics, and Output various Excel spreadsheets for secondary editing.

2.2.2 Development of warehouse management system

Warehouse management system, in fact, is to digitize all cargo information, storage information, transfer information, etc., write to the database, and operate and manage the database. Based on this concept, the management and optimization of the database directly affect the warehouse management system. The development of the database directly determines the development of the warehouse management system. There are various types of existing databases: MySQL, Oracle, Mongo DB, and other databases have their own advantages. With the development of the times, the emergence of cloud databases has made the warehouse management system evolution to the cloud Warehouse management [9].

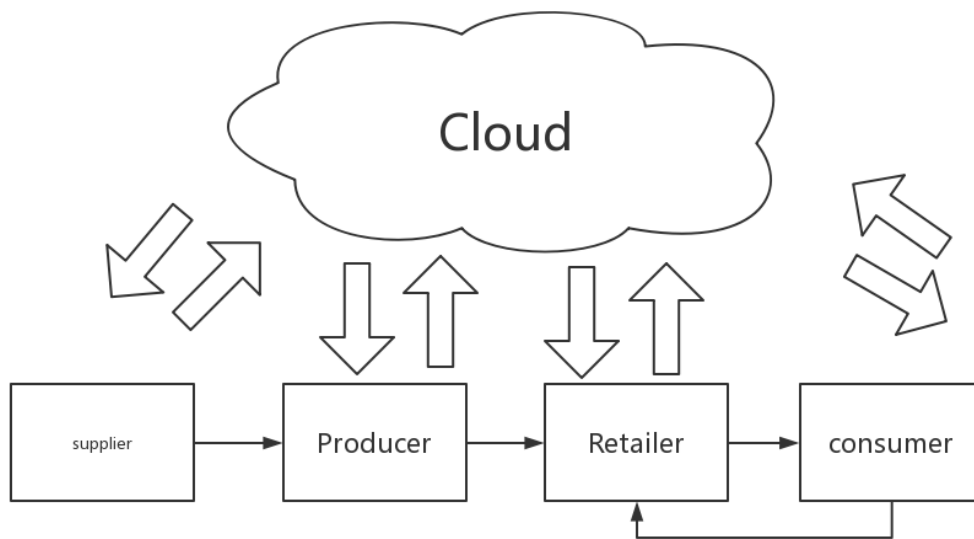


Figure7: Cloud data interaction

Cloud warehousing management, all its data interaction is completed in the cloud, the database structure is not built locally, all services are responsible by the cloud database provider, and users can operate directly in the local, but do not need to set up a local server group to supply the local database Storing data. This is not only a cost saving, but also an improvement in the efficiency of database operations. For the traditional database, there are two common upgrade methods. The first one is to release a new version on the Internet, and the user automatically upgrades. This is generally for individual users, and the second is to send professional personnel to the local database. Upgrades, which are more common in commercial customers and industries where data needs to be tightly protected, such as banks. Every time the database is upgraded and maintained, it is the consumption of human and financial resources. Moreover, in today's society to enter the era of big data, the magnitude of data is no longer comparable to before, which further increases the cost of maintaining the database. The emergence of cloud storage is an upgrade to database operations, and all maintenance and upgrades can be done online in real time in the cloud.

However, cloud services are convenient and contain some hidden dangers. For example, the security, privacy, etc. of storing data in the cloud [10]. In addition, the lack of a standard API prevents customers from easily extracting code and data from the site to run on another site. More generally, outsourcing infrastructure to cloud providers, public cloud customers are bound to face price increases, reliability issues, and even Service provider failures [11].

2.3 Shopping behavior analysis

The analysis of human behavior has already begun very early. As early as in ancient Greece, Aristotle began a certain discussion and observation on human behavior analysis, which is an important beginning for the study of human behavior.[12] The modern analysis of human behavior was proposed by American psychologist Walter Hunter. His view is that psychology has entered the experimental era by what is the debate psychology and what kind of speculative era. Psychology has gone onto the objective the path of human behavior. Shopping behavior analysis is a sub-category of behavior analysis. Through the analysis of people's shopping behavior, we can better understand the selection trend of people when purchasing goods, and thus generate guidance.

2.3.1 The basic framework for shopping behavior analysis

The definition of consumer buying behavior was proposed by Philip Kotler in 2000, which refers to the process activities involved in the search, selection, purchase, use, evaluation, and disposal of products and services to meet needs and desires, including consumers Subjective psychological activities and objective material activities.

Consumers' purchasing motives and purchasing behaviors can be summarized as 6W and 6O, which makes it easier to form a consumer buying behavior framework [13]. These six points are:

- What the market needs (What) - what is the product (Objects).
 - Why buy (Why) - What is the purpose of the purchase (Objectives).
 - Who is the buyer (Who) - What is the organization (Organization)?
 - How to buy (How) - What is the organization's operation behavior (Operation)?
 - When to buy (When) - What is the timing of purchase (Occasions).
 - Where to buy (Where) - What is the purchase (Outlets).
1. By analyzing what consumers want to buy, why do they need it instead of another? Companies can research how to provide the right products to meet the needs of consumers.
 2. By analyzing the formation of purchasing motives (consisting of multiple factors: physiological, economic, social, psychological factors, etc.), understand the purchasing purpose of consumers and adopt corresponding marketing strategies.
 3. Analyze whether the purchaser is an individual, a family or a group, who buys the product for whom, and who is the decision maker, executor, and influence of the

purchase. Based on the analysis, combine the corresponding products, channels, pricing and promotions.

4. Analyze the different requirements of the purchaser for the purchase method, and provide different marketing services in a targeted manner.

5. Analyze the buyer's requirements for the purchase time of a specific product, and timely launch the corresponding products, such as analyzing the impact of natural seasons and traditional holidays on market purchases.

6. Analyze the purchaser's requirements for the purchase location of different products, such as daily necessities, customers generally require the nearest purchase, while the purchases require purchases at the commercial center, while selecting and comparing, special products often require direct purchase to the relevant professional stores, etc. .

2.3.2 Influencing factors of shopping behavior

The factors affecting consumers' purchasing behavior are mainly analyzed from the following two points:

- The first point is the personal impact of consumers.
- The second point is the impact of the store on consumers.

Personal factor

Consumers' purchase behavior will be affected by their personal characteristics, such as gender, age stage, education level, economic income, personality and other aspects will have a certain effect. Take gender as an example: in the traditional physical market, there is a big difference between male and female shopping behavior. For example, men have a lot of rational ingredients when shopping, and often make purchasing decisions after careful consideration, while female consumers have more emotional components when shopping. Often when you see a product you like, you will subconsciously put it into the shopping cart. In addition, male consumers are more autonomous. They often look for information about the price, quality, performance, etc. of their own products, and then make their own judgments; while female online consumers are more dependent when they make shopping. Decision-making often tends to pay more attention to other people's opinions or evaluations. For example, the recommendation of a good friend and the opinions of the family.

Store factor

The influencing factors of the store can be divided into two types, the first is the influence of the style of the store on the customer, and the second is the influence of the display of the product on the customer. The influencing factors of the store can be divided into two types, the first is the influence of the style of the store on the customer,

and the second is the influence of the display of the product on the customer. For the style of the store, traditional brick-and-mortar stores can show their distinctive image through the facade, which attracts consumers. For example, the interior is decorated with bright warm colors to promote people's desire to buy. For the merchandise display, the traditional store can display the merchandise and attract the consumer to purchase through different merchandise display methods. The store can display the beautifully packaged merchandise in a conspicuous place for display, and can also put similar products on the display together.

By analyzing the previous analysis, the consumer's own consumption can be guided. For my project, I can calculate the related product based on the product correlation degree brought by the prior algorithm, and then based on the product. The shelf display plays a guiding role for the customer to adjust the display shelves of the products to guide the consumers to purchase the products. For the goods displayed on the shelves, we can think of it as a kind of state in the library, so that we can achieve the purpose of improving the efficiency of inventory management and increasing the sales volume. Because the customer that the project is facing is a small and small retail industry practitioner, it is unrealistic for the field installation project. Through the cloud storage of the database, the whole project can be said to be a web application. It is in line with the overall expectation, and the small and micro retail industry will not generate a large amount of sales data, and its data possession will not have too much impact on the a priori algorithm, so the construction of the project is completely feasible.

Chapter III

System Design

3.1 Overview of tools and technology

The tools and technologies used in the construction of this project include the following:

- Front-end: CSS + bootstrap + JavaScript + HTML5 + jQuery
- Back-end: J2EE+JAVA+eclipse+Maven
- Database: MySQL

3.1.1 Front end

For a web application, the front-end quality directly affects the system evaluation of the entire project. The display effect of the page is very important for a web application. While meeting the user's needs, it also needs to take into account the beauty of the page. In terms of the habits of user operations, etc., I used the related technology of the current build page to build the overall page accordingly.

1. CSS(Cascading Style Sheets) :

CSS is a computer language used to add structural styles (fonts, spacing, colors, etc.) to structured documents such as HTML documents or XML applications. CSS can't be used alone, it must work with HTML or XML to decorate HTML or XML. The most important goal of CSS is to separate the contents of the file from its display. [14] In the initial web page production, because there is no CSS, all the HTML files contain information displayed by the file, such as font color, page background, image format, etc., which need to be listed in the HTML page in turn, resulting in the overall The code redundancy of the page is high and the cost of page maintenance is high. The advent of CSS allows authors to isolate most of this information and simplify HTML files, which are placed in an auxiliary file written in CSS. The HTML file contains only the structure and content information, and the CSS file contains only the style information. This way the overall page is a lot of neat and tidy, you only need to modify the corresponding

CSS file when you need to modify the display style.

CSS consists of multiple rules, each consisting of selector, property, and value. [14] The selectors are mainly composed of five types, namely tag selector, class selector, ID selector, universal selector, and attribute selector, each of which is used in a different manner. Figure 1 shows how the tag selector, class selector, and ID selector are used.

```
h1 {  
    font-size: 3.5em;  
    color: #fff;  
    letter-spacing: 3px;  
    text-align: center;  
    font-style: italic;  
    margin: 3vw 1vw;  
    font-family: 'Josefin Sans', sans-serif;  
}  
  
.ft02 {  
    color: #4E4C7A  
}  
  
#datediv {  
    height: 100px;  
    width: 630px;  
    overflow: hidden;  
    margin: 40px auto;  
    text-align: center;  
    margin: auto;  
}
```

Figure3-1: CSS selector call mode

2. HTML5

HTML5 is the latest version of HTML, completed by the World Wide Web Consortium (W3C) in October 2014 [15]. The goal is to replace the previously customized legacy HTML standard and to provide a service standard that meets the needs of high-speed Internet in the modern era of rapid Internet development. HTML5 is a set of combined technologies including CSS, JAVASCRIPT, and HTML. HTML5 adds a lot of new element tags, such as <video>, <audio>, <canvas>, etc., to better support the addition and processing of multimedia and image content in web pages. At the same time, some labels under the old standard have been modified and redefined to make it more standardized.

3. JQUERY

JQuery is a library of JavaScript supporting browsers. The purpose is to simplify the operation between the browser and JavaScript. 73.9% of the top 10,000 most visited websites in the world use jQuery, which is currently the most popular JavaScript function library.[16] JQuery is an open source software. JQuery's syntax design makes many operations simple, and it can do a lot of operations in a more streamlined language, such as manipulating document objects, creating animation effects, handling events, and creating Ajax programs.

Creating Ajax program is a major feature of JQuery. It is also a program I use more in this project. Ajax is called Asynchronous JavaScript and XML. In short, when we need to request information from the same page, the page does not need to refresh the load data but "asynchronous" execution load, for example, when we purchase the ticket, after selecting the destination, the page automatically displays the destination-related ticket without changing or overwriting the original selection condition, which is a case of asynchronously refreshing a page. And this feature makes Web applications respond more quickly to user actions and avoids sending unchanging messages over the network.

4. JAVASCRIPT

JavaScript [17], often abbreviated as JS, is an advanced, direct language programming language. It supports object-oriented programming, imperative programming, and functional programming. It provides syntax for manipulating text, arrays, dates, regular expressions, and more. And it is used by most of the world's websites, and is also supported by the world's major browsers (Chrome, IE, Firefox, Safari, and Opera). The language design of JavaScript is mainly influenced by Self (a prototype-based programming language) and Scheme (a functional programming language) [18]. In terms of grammatical structure, it is similar to C language (such as if conditional statement, while loop, switch statement, do-while loop, etc.) [19].

JS is different from PHP's server-side language. JavaScript runs primarily as a client-side scripting language on the user's browser, without server support. This feature allows JavaScript to reduce the burden on the server side, and JavaScript also supports the cross-platform operation. In JS, variables are defined using the 'var' keyword. JS is also like most scripting languages types are associated with values rather than with variables. For example, x variables can be numeric and can be assigned to strings.

5. BOOSTRAP

Bootstrap [20] is a set of open source front-end frameworks for Web and Web application development that was officially released as an open source project in 2011. It includes HTML, CSS and JavaScript frameworks that provide typography, forms, buttons, navigation and various other components and JavaScript extensions. Its purpose is to make the development of web applications easier. And now most browsers support Bootstrap.

Bootstrap's modular design is essentially CSS-compliant, defines a basic set of HTML components, and creates a unique contemporary style for text, tables, and form elements. It also supports JavaScript via jQuery, providing dialogs and tooltips. It also enhances the functionality of some user interface elements.

3.1.2 Back end

The background part mainly contains the transaction processing method. This part is the core of the whole project. All the running logic and the execution order of the program are completed in this part. At the same time, this part also contains the data interaction with the database. It is the whole project. Most important part.

1. Java

Java [21] is a computer high-level programming language that was first developed by Sun and widely used in various fields. It is open source, has a cross-platform, object-oriented design and generic programming features⁶. It is widely used in enterprise-class Web. Application development and mobile application development. The style of the Java programming language is very close to the C++ language. Inheriting the core of C++ language object-oriented technology, abandoning the pointers that are easy to cause errors in C++, replacing by reference; removing the operator overloading and multiple inheritance features in C++, replacing with interfaces; and also adding garbage collector function. A major feature of the Java language is that programs written in Java have good support for cross-platform operations. This is because the Java language is different from other compiled languages. When compiling, Java first compiles the source code into a byte code file. After that, the implementation of the byte code file is

⁶ Is a style or paradigm of programming language? Generics allow programmers to use code that is specified later when writing code in a strongly typed programming language, specifying them as arguments when instantiating.

interpreted according to the compiler of different platforms, thus having such a good cross-platform capability.

There are also things to note when writing Java programs:

Case sensitive: Java is case sensitive, which means that identifiers 'A' and 'a' are different.

Class name: For all classes, the first letter of the class name should be capitalized. If the class name consists of several words, the first letter of each word should be capitalized.

Method name: All method names should start with a lowercase letter. If the method name contains several words, the first letter of each subsequent word is capitalized.

Source file name: The source file name must be the same as the class name. When saving a file, you should use the class name as the file name (consistent with the first feature). If the file name and class name are not the same, it will cause a compilation error.

2. J2EE

The full name of J2EE is called Java Platform Enterprise Edition. It is the standard platform launched by Sun for Java enterprise application development. In March 2018, it was renamed Jakarta EE. There are many technical standards in the J2EE platform. For example, for this project, the standards used are: JDBC(Java Database Connectivity), JMS(Java Message Service), WS(Web Services), Servlet(Java Servlet API), JNDI(Java Naming and Directory Interface),JSP (Java Server Pages) etc.

3. Eclipse

Eclipse is a well-known cross-platform open source integrated development environment (IDE). Originally used primarily for Java language development, although it is just a framework platform, the support of many plug-ins makes Eclipse more flexible. The plugin mechanism of Eclipse is lightweight software architecture. That is, all plug-ins can be downloaded and installed freely by the user, and the plug-in provides all the additional functions, such as support for languages other than Java.

4. Maven

Apache Maven [22] is a software project management and comprehension tool. Based on the concept of a project object model (POM), Maven can manage a project's build,

reporting and documentation from a central piece of information. Figure 2 shows the engineering structure of the maven project.

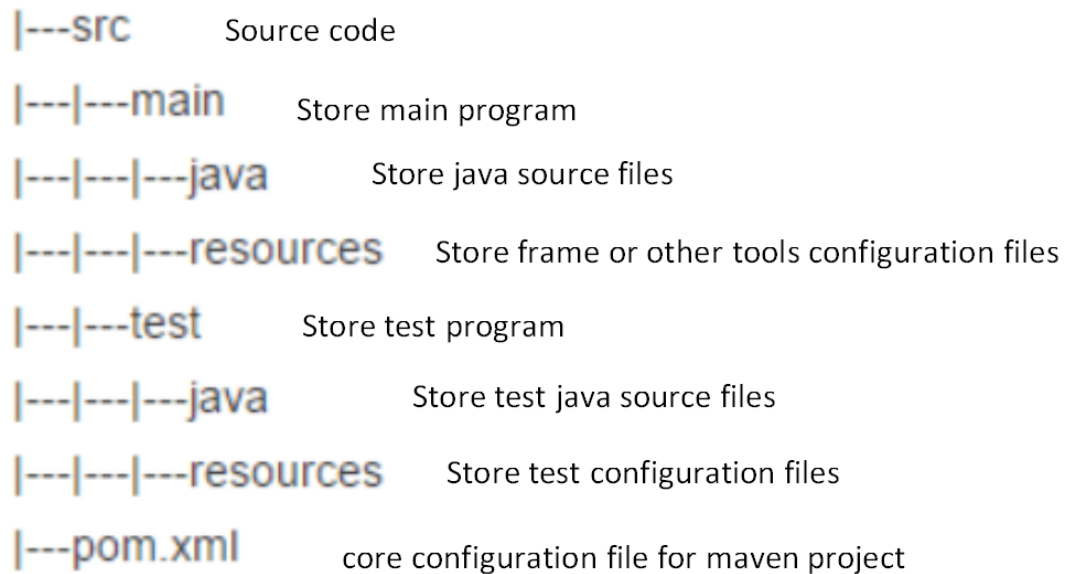


Figure3-2: Maven project structure

Why maven has a structure like this because the first maven responds for project automatic build, such as compile, this structure can help maven find java source files location, Don't need to manually specify the location. Secondly, this structure not only reduces labor but also prevents mistakes.

One of the great features of the maven project is the Dependent library management. In the maven project, we do not need to manually import the jar package as needed for the general java project. It only needs to be added and modified in the pom file. In the maven project, it imports the jar package by coordinates, which are group Id, Artifact Id, and version. In Figure3, I used JUnit jar package, so you can see in dependency label have sub-labels which group Id is JUnit, artifact Id is JUnit and version is 4.0. After the pom file is configured maven will download this package form Central repository, then we can use this jar package. If this jar package has a relationship with other jar packages, then maven will automatically import the jar package with dependencies and do not need to import it manually.


```

<dependencies>
  <dependency>
    <groupId>junit</groupId>
    <artifactId>junit</artifactId>
    <version>4.0</version>
    <scope>test</scope>
  </dependency>
</dependencies>

```

Figure3-3: maven dependent library management

3.1.3 Database side

The database side is the most important place in the project, because this is the source of all the page display content, and is the basis for all operations. For this project, the data processing on the database side directly affects the project display.

1. MySQL

MySQL is an open relational database management system (RDBMS). MySQL has become the most popular open source database due to its high performance, low cost, and high reliability, so it is widely used in small and medium-sized websites on the Internet. MySQL has many features, such as MySQL supports multitasking, makes full use of CPU resources, and supports multiple users. And MYSQL optimizes the SQL query algorithm, which effectively improves the query speed. It also supports multiple database connection paths such as TCP / IP, ODBC and JDBC. Most importantly, it also supports the creation and maintenance of stored procedures and views.

Stored Procedure is a set of SQL statements in a large database system that is used to complete a specific function. It is stored in the database. After the first compilation, the call does not need to be compiled again. The user specifies the name of the stored procedure and gives it Execute the parameter (if the stored procedure has parameters) to execute it. A stored procedure is an important object in the database.

A view is a virtual table that is exported from one or several basic tables (or views). A view is a transformation of raw database data and is another way to view data in a table. You can think of a view like a moving window through which you can see the data of interest. Views are obtained from one or more actual tables whose data is stored in the database. The tables used to generate the view are called the base table of the view. A view can also be generated from another view.

The use of stored procedures and views is a very important aspect of data manipulation on the database side of the project. It is convenient for the background program to simply find the required data and data structure, and do not perform secondary processing.

3.2 System architecture

3.2.1 Architecture

Basic on the topic is web application so that to the structural components that should be are following:

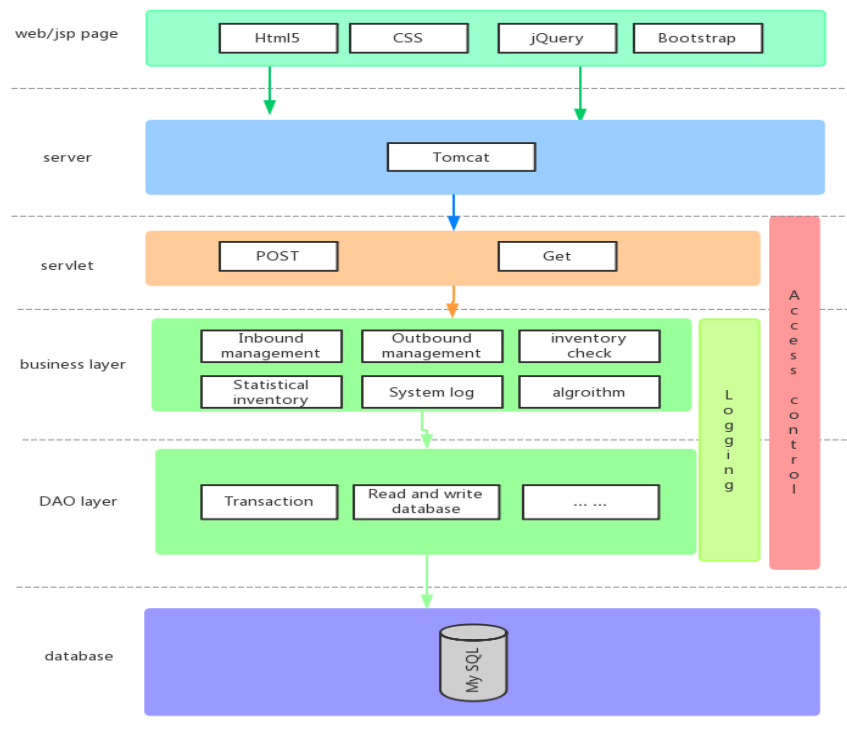


Figure 3-4: System architecture

3.2.2 Execution logic (data flow chart)

The program execution logic graphic is following:

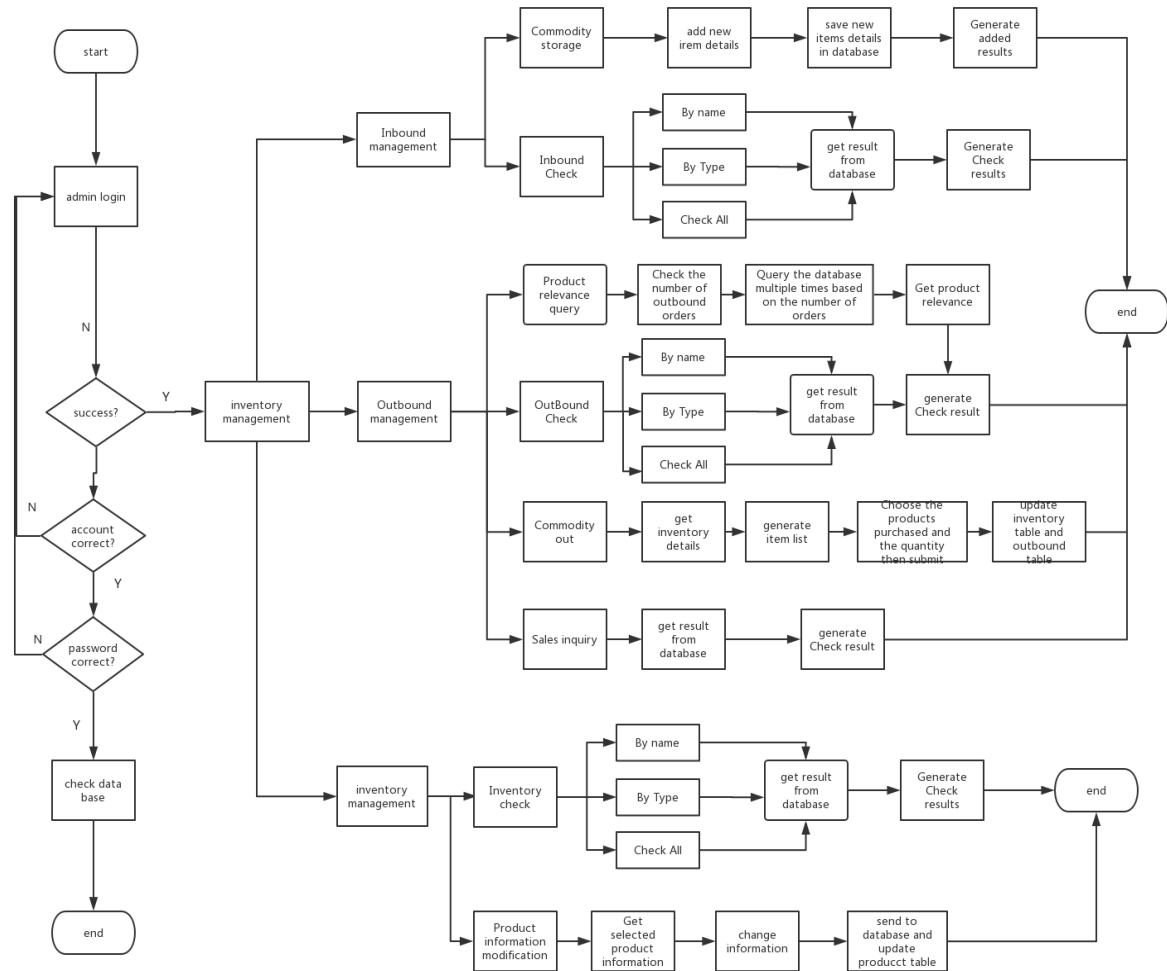


Figure 3-5 Program execution logic diagram

3.2.3 Database E-R diagram

The database E-R diagram is following:

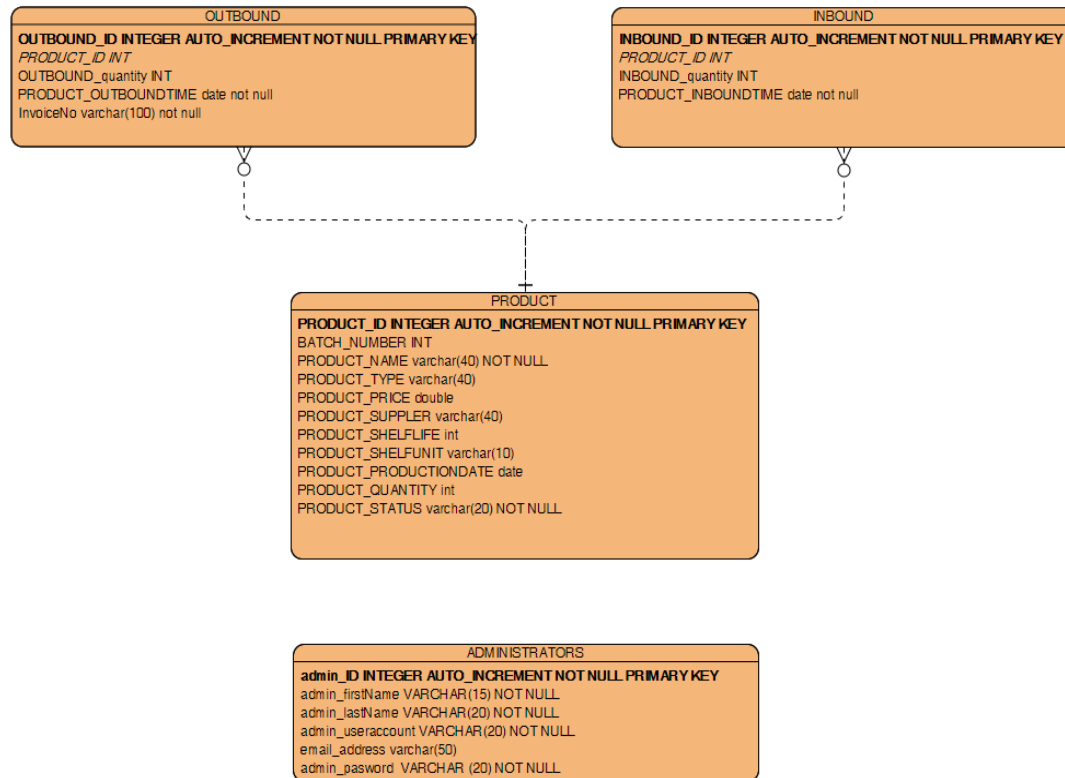


Figure 3-6 database E-R diagram

For the interpretation of the ER diagram, the project mainly realizes the intelligent management of commodity inventory, so the design of the database is mainly considered from the aspects of commodity storage (INBOUND), commodity delivery (OUTBOUND), and commodities (PRODUCTS). For a retail store, the product is its only meaningful data entity, so when designing the database, everything is centered on the commodity entity. A product can be put into the warehouse multiple times, or it can be purchased multiple times. Therefore, the product table has a one-to-many relationship with the inbound table and the outbound table. At the same time, in the database, many views and stored procedures are created to help the data display better. The user table (ADMINISTRATORS) is designed to control user login, and the user can choose to log in to the management system based on the account or email.

3.2.4 UML sequence diagram

For the web application, the execution order is basically the same. The first step is to operate the user in the browser, send the request, the background code accepts the request, processes the request, goes to the database to query related data or perform other operations. After that, the database returns the query result to the background code, and the background code then processes the data back to the page, and the page parses the data and finally displays the result.

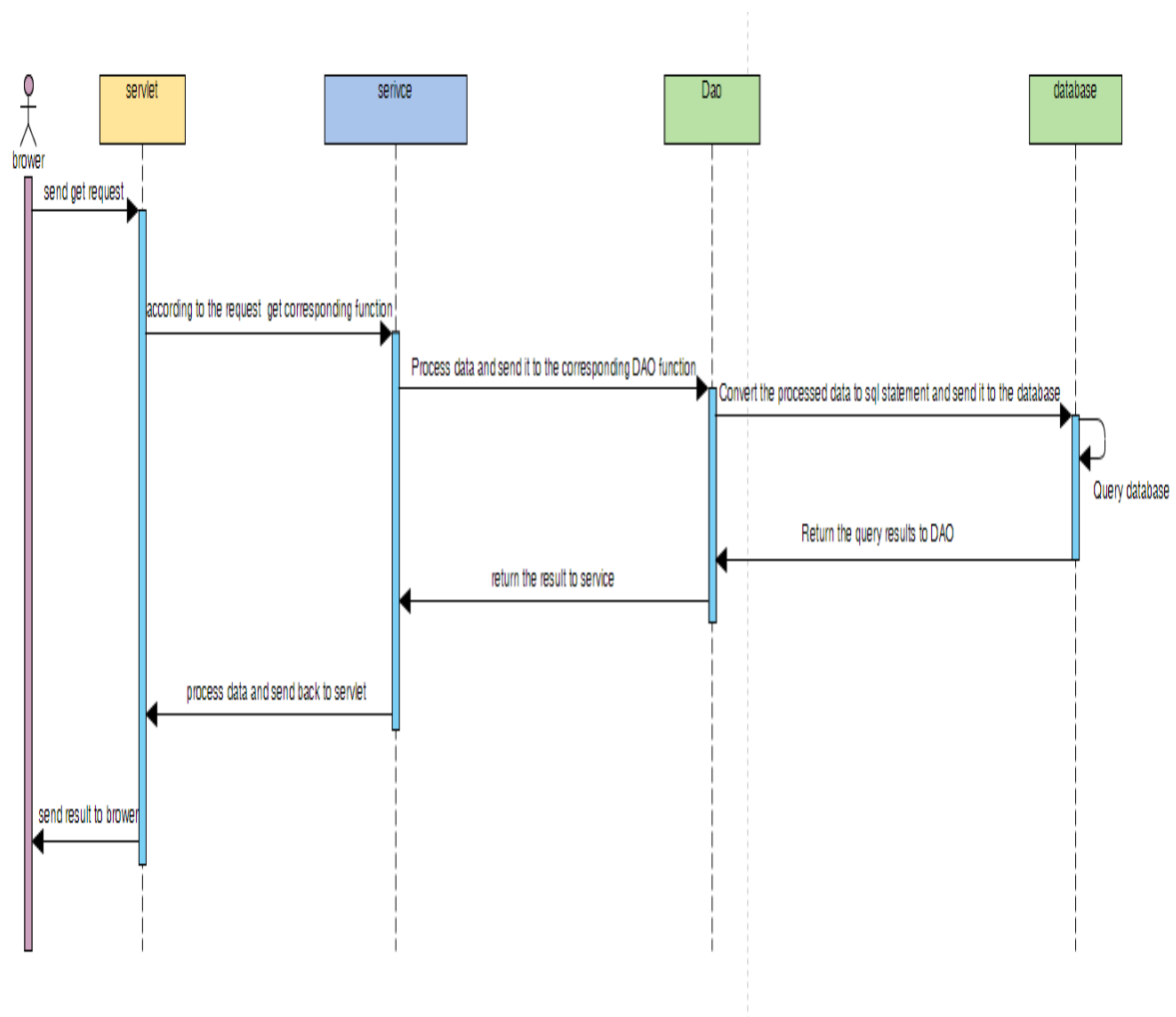


Figure 3-7 sequence diagram

3.2.5 Use case diagram

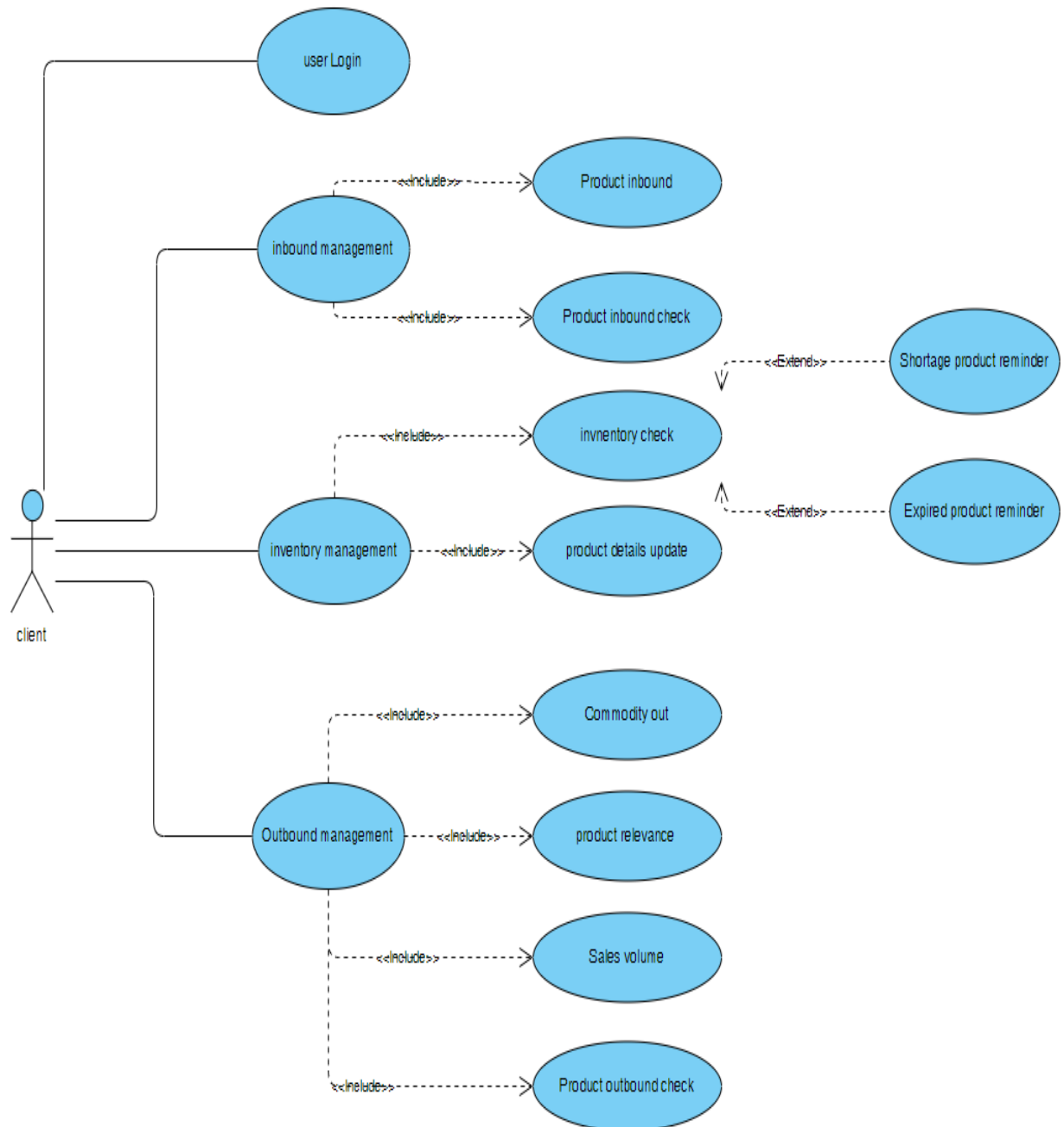


Figure 3-8 Use case diagram

3.3 Requirement

According to the use case diagram and the execution logic of the system, the requirements of the system are the following:

Requirement1	User Login	
Goal in Context	User can login to the system and start to use it.	
Preconditions	Establish an internet connection	
Success Post Condition	Web page display the Login page	
Failed Post Condition	Browser cannot load the login page and return 404 page code	
Actors	Admin	
Trigger	User access the web application	
DESCRIPTION	Step	Action
	1	The user access the web application
	2	Browser display the login page
	3	User fill in username or email and password
	4	System cannot link to dashboard
SUB-VARIATIONS		
		Branching Action
	1	User closes web application
	2	Lost network connection

Requirement 2	Product inbound	
Goal in Context	Users can add new product into system	
Scope & Level	Inbound management & Primary Task	
Preconditions	After user success login	
Success Post Condition	Add new product details into the system	
Failed Post Condition	The system did not add new product information.	
Actors	Admin	
Trigger	User click on commodity storage in product inbound management	
DESCRIPTION	Step	Action
	1	The user clicks on the commodity storage tag in the navigation bar.
	2	System loading new product page information
	3	User fills in the new product information
	4	After the user fills in, click the OK button.
SUB-VARIATIONS		Branching Action
	1	User closes web application
	2	User filled out information that does not meet the specifications
	3	Lost network connection

Requirement3	Product inbound check	
Goal in Context	Users can view the inbound details of existing stock product	
Scope & Level	Inbound management Primary Task	
Preconditions	After user success login	
Success Post Condition	Generate product inbound information table	
Failed Post Condition	Not generate the product inbound information table	
Actors	Admin	
Trigger	User click on inbound check tag inside the inbound management	
DESCRIPTION	Step	Action
	1	The user clicks on the inbound check
	2	System loading page information
	3	User selects the query mode (by name, by type, get all information)
	4	After select mode, click on OK button.
SUB-VARIATIONS		Branching Action
	1	User closes web application
	2	Lost network connection

Requirement4	Product inventory check	
Goal in Context	Users can check the details and status of products already in stock	
Scope & Level	Inventory management & Primary Task	
Preconditions	After user success login	
Success Post Condition	Generate product inventory information table	
Failed Post Condition	Not generate the product inventory information table	
Actors	Admin	
Trigger	User click on inventory check	
DESCRIPTION	Step	Action
	1	The user clicks on the inventory check in the navigation bar.
	2	System loading inventory check page information
	3	User selects the query mode (by name, by type, get all information)
	4	After select mode, click on OK button.
SUB-VARIATIONS		Branching Action
	1	User closes web application
	2	Lost network connection

Requirement5	Storage product reminder	
Goal in Context	Users can check which products are insufficient and need to replenish	
Scope & Level	Inventory management & Primary Task	
Preconditions	After user success login	
Success Post Condition	System automatically generate inventory shortage table with an inventory quantity less than 100	
Failed Post Condition	Not automatically generate the inventory shortage table	
Actors	Admin	
Trigger	User click on inventory check tag inside the inventory management	
DESCRIPTION	Step	Action
	1	The user clicks on the inventory check in the navigation bar.
	2	System loading product check page information
	3	Automatically generate the inventory shortage table
SUB-VARIATIONS		Branching Action
	1	User closes web application
	2	Lost network connection

Requirement6	inventory expired product reminder	
Goal in Context	Users can see which products have insufficient sales days to develop a corresponding promotion strategy.	
Scope & Level	Inventory management & Primary Task	
Preconditions	After user success login	
Success Post Condition	The system automatically generates a list of products with a shelf life of less than one month.	
Failed Post Condition	Not automatically generate the expired product table	
Actors	Admin	
Trigger	User click on inventory check tag inside the inventory management	
DESCRIPTION	Step	Action
	1	The user clicks on the inbound check in the navigation bar.
	2	System loading product check page information
	3	automatically generate the expired product table
SUB-VARIATIONS		Branching Action
	1	User closes web application
	2	Lost network connection

Requirement7	Product information update	
Goal in Context	Users can update product information	
Scope & Level	Inventory management & Primary Task	
Preconditions	After user success login and successful generate the product information table	
Success Post Condition	Product details change successful.	
Failed Post Condition	Product details not changed	
Actors	Admin	
Trigger	User click on edict link inside the product information table	
DESCRIPTION	Step	Action
	1	The user clicks on the edict link inside the product information table.
	2	System pops up a dialog box containing the selected product details.
	3	User updates the information they want to change
	4	After update, click on OK button.
SUB-VARIATIONS		Branching Action
	1	User closes web application
	2	Lost network connection

Requirement8	Commodity out	
Goal in Context	Users can use this program to sell products	
Scope & Level	Outbound management & Primary Task	
Preconditions	After user success login	
Success Post Condition	Automatically generate product information table and can add selected product into shopping cart.	
Failed Post Condition	Not automatically generate the product information table	
Actors	Admin	
Trigger	User click on Commodity out tag inside the outbound management	
DESCRIPTION	Step	Action
	1	The user clicks on the commodity out inside the outbound management in the navigation bar.
	2	System loading commodity out page information
	3	User selects the product they want and click buy tag inside the product table
	4	Fill in the quantity to be purchased and click OK
SUB-VARIATIONS		Branching Action
	1	User closes web application
	2	Lost network connection

Requirement9	shopping cart for Commodity out	
Goal in Context	Users can add the products they want to buy to the shopping cart.	
Scope & Level	outbound management & Primary Task	
Preconditions	User success login	
Success Post Condition	Generate the shopping cart and add new information about the product which costumer want to buy and calculate the price	
Failed Post Condition	Not generate the shopping cart	
Actors	Admin	
Trigger	After the user decides the purchase quantity and clicks the OK button	
DESCRIPTION	Step	Action
	1	User select the product and click the buy link inside the product table
	2	User decides to purchase the quantity and clicks OK
SUB-VARIATIONS		Branching Action
	1	User closes web application
	2	Lost network connection

Requirement10	Product relevance	
Goal in Context	Users can see the degree of association between the products sold, so that the shelves can be arranged more reasonably.	
Scope & Level	Outbound management & Primary Task	
Preconditions	After user success login	
Success Post Condition	Automatically generate Product association table	
Failed Post Condition	Not automatically generate Product association	
Actors	Admin	
Trigger	User click on product relevance tag inside the outbound management	
DESCRIPTION	Step	Action
	1	User click on product relevance tag inside the outbound management
	2	System loading product relevance page information
SUB-VARIATIONS		Branching Action
	1	User closes web application
	2	The amount of data sold in the product does not meet the minimum amount of data required by the algorithm
	3	Lost network connection

Requirement11	Product sales ranking	
Goal in Context	Users can view their store sales to find out which product is more popular	
Scope & Level	Outbound management & Primary Task	
Preconditions	After user success login	
Success Post Condition	Automatically generate sales ranking table	
Failed Post Condition	Not automatically generate sales ranking table	
Actors	Admin	
Trigger	User click on product relevance tag inside the outbound management	
DESCRIPTION	Step	Action
	1	User click on product relevance tag inside the outbound management
	2	System loading product relevance page information
SUB-VARIATIONS		Branching Action
	1	User closes web application
	2	No products have been sold in the system yet
	3	Lost network connection

Requirement12	Product outbound check	
Goal in Context	Users can view the outbound details of product	
Scope & Level	Outbound management & Primary Task	
Preconditions	After user success login	
Success Post Condition	Generate product outbound information table	
Failed Post Condition	Not generate the product outbound information table	
Actors	Admin	
Trigger	User click on outbound check tag inside the outbound management	
DESCRIPTION	Step	Action
	1	The user clicks on the outbound check in the navigation bar
	2	System loading product out bound check page information
	3	User selects the query mode (by name, by type, get all information)
	4	After select mode, click on OK button.
SUB-VARIATIONS		Branching Action
	1	User closes web application
	2	Lost network connection

Character IV

System implementation and Testing

4.1 Implementation

According to the use case diagram, the system execution logic and requirements, can be derived from the specific system implementation as follows:

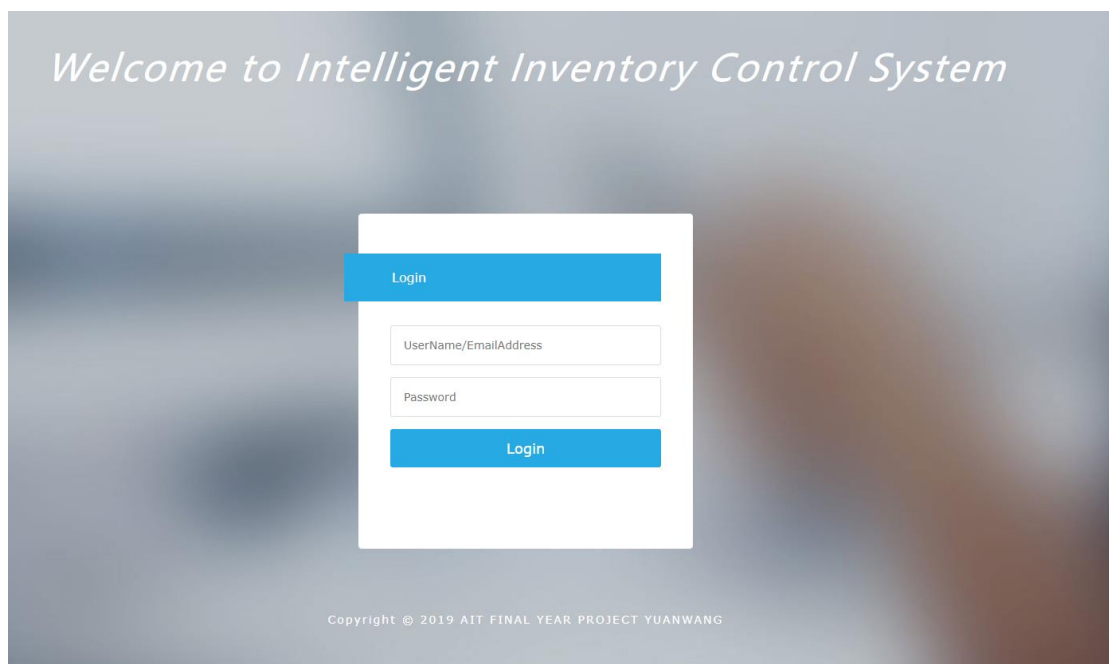


Figure 4-1: user login page

This is the user login interface of the system.

This interface allows the user to log in with an account or a mailbox. The specific implementation uses CSS+HTML5+JQUERY+JS in the front end. The back-end code received the parameter from the front end through the HTTP GET request.

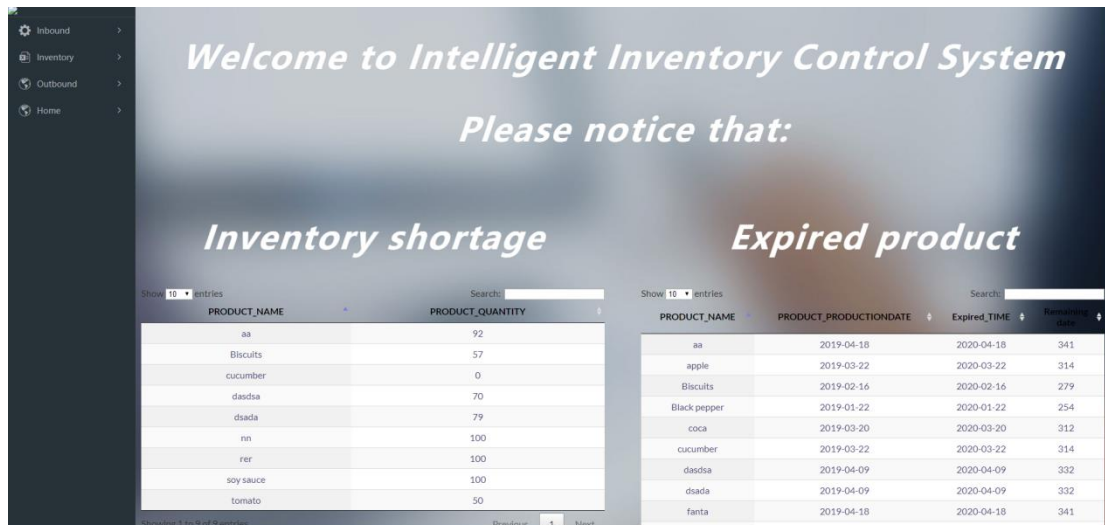


Figure 4-2 dashboard

This is the main interface of the system. The system navigation bar is located on the left side. It is implemented by JQuery. The left and right layout of the overall display page is implemented by the HTML ‘frameset’ framework. The right side is the main data display page. The two tables displayed are implemented by DataTable function in jquery, text style, background pattern and overall page display are implemented by CSS + JQuery.

new product details

BATCH_NUMBER:

product name:

type:

price:

supplier:

quantity:

Production Date:

Shelf life:

Shelf life unit:

status:

Close

save

Figure 4-3 add new item dialog box



The screenshot shows a web interface titled "Commodity storage Result". At the top, there is a search bar and a dropdown menu showing "10 entries". Below this is a table with the following columns: PRODUCT_NAME, BATCH_NUMBER, PRODUCT_TYPE, PRODUCT_PRICE, PRODUCT_SUPPLIER, PRODUCT_QUANTITY, PRODUCT_PRODUCTIONDATE, SHELF_LIFE, and SHELF_LIFE_UNIT. The table contains one row of data: "testcase", "17", "drink", "2.58", "Yuan", "100", "2019-04-25", "12", and "month". At the bottom of the table, it says "Showing 1 to 1 of 1 entries". On the right side, there are "Previous" and "Next" buttons, with "1" in the middle.

PRODUCT_NAME	BATCH_NUMBER	PRODUCT_TYPE	PRODUCT_PRICE	PRODUCT_SUPPLIER	PRODUCT_QUANTITY	PRODUCT_PRODUCTIONDATE	SHELF_LIFE	SHELF_LIFE_UNIT
testcase	17	drink	2.58	Yuan	100	2019-04-25	12	month

Figure 4-4 add new item result table

Figures 4-3 and 4-4 show the interface for adding product information.

Figure 4-3 shows the dialog box that pops up when adding product information. It is implemented by the combination of bootstrap and jQuery, in which the text field the added verification is implemented by JS + HTML5.

Figure 4-4 shows the results returned after adding a new product. Whenever the return result table appears, it means that add new product information is successful. In the front end, the function is displayed by JQuery. In the back end, the data transfer form is transmitted by AJAX and use a Servlet to receive it. The main method used is HTTP GET request. The main data format is JSON.



The screenshot shows a form titled "Please choice options". There are three buttons: "Type", "Name", and "ALL". Below the buttons is a text input field with the label "name:" and the value "Waffles". At the bottom, there is a "Submit" button.

Please choice options

Type Name ALL

name: Waffles

Submit

Figure 4-5 inbound check mode select

Please choice options

Type Name ALL

name: Waffles

Submit

Check Result

product_name	product_type	batch_number	product_supplier	inbound_time	inbound_quantity
Waffles	Snack	1	supplier001	2019-03-24	5000

Showing 1 to 1 of 1 entries

Previous 1 Next

Figure 4-6 inbound check result

Figure 4-5 and Figure 4-6 show the implementation of the inbound Check.

4-5 shows the query mode selection, the dynamic performance of the button selection is implemented by jQuery, the relevant text field after clicking the button the display and hiding are implemented by JQuery in conjunction with CSS. The main transmission method is implemented by JavaScript + JQuery + Ajax. The back-end code obtains the keyword to query the data by obtaining the HTTP GET request sent by Ajax. Its main data delivery format is JSON.

Figure 4-6 shows the results of the query, implemented by JQuery.

Please choice options

Type Name All

name: Biscuits

Submit

Figure 4-7 inventory check mode select

Check Result

PRODUCT_NAME	PRODUCT_TYPE	PRODUCT_PRICE	PRODUCT_SUPPLIER	PRODUCT_PRODUCTIONDATE	PRODUCT_QUANTITY	PRODUCT_STATUS	edit
Biscuits	Snack	1.09	Oreo	2019-02-16	57	ok	• Edit

Showing 1 to 1 of 1 entries

Inventory shortage

PRODUCT_NAME	PRODUCT_QUANTITY
aa	92
Biscuits	57
cucumber	0
daada	70
daada	79
nn	100
rer	100
soy sauce	100
testcase	100
tomato	50

Showing 1 to 10 of 10 entries

Expired product

PRODUCT_NAME	PRODUCT_PRODUCTIONDATE	Expired_TIME	Remaining date
aa	2019-04-18	2020-04-18	341
apple	2019-03-22	2020-03-22	314
Biscuits	2019-02-16	2020-02-16	279
Black pepper	2019-01-22	2020-01-22	254
coca	2019-03-20	2020-03-20	312
cucumber	2019-03-22	2020-03-22	314
daada	2019-04-09	2020-04-09	332
daada	2019-04-09	2020-04-09	332
fanta	2019-04-18	2020-04-18	341
fanta	2019-04-10	2020-04-10	333

Showing 1 to 10 of 23 entries

Figure 4-8 inventory check result

Figures 4-7 and 4-8 show the inventory check, and 4-7 show the query mode selection, which is implemented in the same way as the product inbound check. 4-8 shows the inventory check result, the stock shortage table, and the expired product table. These three result tables are all implemented by the JQuery DataTable function. On the back end, use a Servlet to receive HTTP GET requests sent by Ajax.

product information ×

product name:

type:

price:

supplier:

quantity:

Production Date:

Shelf life:

Shelf life unit:

status:

Figure 4-9 inventory product update

Check Result

Show 10 entries

PRODUCT_NAME	PRODUCT_TYPE	PRODUCT_PRICE	PRODUCT_SUPPLIER	PRODUCT_PRODUCTIONDATE	PRODUCT_QUANTITY	PRODUCT_STATUS	edit
Biscuits	Snack	1.09	OreoOOO	2019-02-18	57	ok	Edit

Showing 1 to 1 of 1 entries

Previous 1

Inventory shortage *Expired product*

Figure 4-10 inventory product update result

Figure 4-9 and Figure 4-10 show the product information update. When the inventory check is completed, click the last edit link in the row to display the dialog box shown in Figure 4-9. The pop-up window is implemented using bootstrap. The limitation of the text field is implemented by html5+regular expression. The data transmission is implemented by JavaScript + Ajax. The back-end Servlet code receives the HTTP POST request sent by Ajax to obtain the data. The data transmission format is JSON. Figures 4-10 show the results of the modified data, which have the same implementation shown in 4-8.

Result

Show 10 entries

product name	Association rule
Ham sausage,->Instant noodles,	0.75
Ham sausage,moon cake,->Instant noodles,	0.8571428571428571
Instant noodles,moon cake,->Laoganma,	0.8
Laoganma,Instant noodles,->moon cake,	0.8888888888888888
Laoganma,moon cake,->Instant noodles,	0.8888888888888888
moon cake,->Instant noodles,	0.8333333333333334
moon cake,->Laoganma,	0.75
orange,->Instant noodles,	0.7
soy sauce,->Instant noodles,	0.7272727272727273
Waffles,->Instant noodles,	1.0

Showing 1 to 10 of 10 entries

Previous 1 Next

selling information

Show 10 entries

product name	Sales
aa	18
apple	290
Biscuits	60
Black pepper	39
coca	405
cucumber	51
daidsa	30
dda	53
dsada	121
fanita	169

Showing 1 to 10 of 22 entries

Previous 1 Next

Figure 4-11 inventory product relevance and sales ranking result


```

14
15 public class Apriori {
16     private final static int SUPPORT = 6; // minimum support
17     private final static double CONFIDENCE = 0.7; // minimum confidence
18
19     private final static String ITEM_SPLIT = ","; // Separator between items
20     private final static String CON = "-->"; //Separator between items
21
22
23
24     //Initialize the transaction record.
25     //In the apriori algorithm, should ensure that the items in the item set are ordered.
26
27     public Map<String, Integer> getFC(ArrayList<String> dataList) {
28         //Store all frequent sets
29         Map<String, Integer> frequentCollectionMap = new HashMap<String, Integer>();
30
31         frequentCollectionMap.putAll(getItem1FC(dataList));
32
33         Map<String, Integer> itemkFcMap = new HashMap<String, Integer>();
34         itemkFcMap.putAll(getItem1FC(dataList));
35         while (itemkFcMap != null && itemkFcMap.size() != 0) {
36             Map<String, Integer> candidateCollection = getCandidateCollection(itemkFcMap);
37             Set<String> ccKeySet = candidateCollection.keySet();
38             //Cumulative count of candidate set items
39             for (String trans : dataList) {
40                 for (String candidate : ccKeySet) {
41                     boolean flag = true;
42                     //Used to determine whether the candidate appears in the transaction.
43                     //If it appears, the count is incremented by 1.
44

```

Figure 4-12 Apriori Algorithm code segment

```

44
45         String[] candidateItems = candidate.split(ITEM_SPLIT);
46         for (String candidateItem : candidateItems) {
47             if (trans.indexOf(candidateItem + ITEM_SPLIT) == -1) {
48                 flag = false;
49                 break;
50             }
51         }
52         if (flag) {
53             Integer count = candidateCollection.get(candidate);
54             candidateCollection.put(candidate, count + 1);
55         }
56     }
57 }
58 //Get frequent set items that meet the support level from the candidate set
59 itemkFcMap.clear();
60 for (String candidate : ccKeySet) {
61     Integer count = candidateCollection.get(candidate);
62     if (count >= SUPPORT) {
63         itemkFcMap.put(candidate, count);
64     }
65 }
66 //Merge all frequent sets
67
68 frequentCollectionMap.putAll(itemkFcMap);
69

```

Figure 4-13 Apriori Algorithm code segment

```

139 private Map<String, Integer> getItem1FC(ArrayList<String> dataList) {
140     Map<String, Integer> sItem1FcMap = new HashMap<String, Integer>();
141     Map<String, Integer> rItem1FcMap = new HashMap<String, Integer>();
142
143     for (String trans : dataList) {
144         String[] items = trans.split(ITEM_SPLIT);
145         for (String item : items) {
146             Integer count = sItem1FcMap.get(item + ITEM_SPLIT);
147             if (count == null) {
148                 sItem1FcMap.put(item + ITEM_SPLIT, 1);
149             } else {
150                 sItem1FcMap.put(item + ITEM_SPLIT, count + 1);
151             }
152         }
153     }
154
155     Set<String> keySet = sItem1FcMap.keySet();
156     for (String key : keySet) {
157         Integer count = sItem1FcMap.get(key);
158         if (count >= SUPPORT) {
159             rItem1FcMap.put(key, count);
160         }
161     }
162     return rItem1FcMap;
163 }
164

```

Figure 4-14 Apriori Algorithm code segment

Figure 4-11 shows the relevance of the product and the sales ranking. The left side of the figure shows the relevance of the product. The acquisition of the correlation is based on the results generated by the Apriori algorithm. The minimum support of the Apriori algorithm is 6, and the minimum confidence is 0.7. According to the results, it can be seen that the association degree between the displayed products is greater than 0.7, indicating that the algorithm works normally. The data used is obtained through the database and in the database side, the view is used in order to better return the required data and structure. The right side of the figure shows the sales ranking. The two result display tables are implemented by jQuery + JS + Ajax. The back-end receives the HTTP GET request sent by the front end to find the corresponding function to process the data.

Figure 4-12, 4-13, and 4-14 show the segment of the Apriori algorithm code, in the 4-12 the code snippet, static variables A and B in lines 16 and 17 define the minimum support and minimum confidence required by the algorithm, which is the two most important thresholds we use for a priori algorithms. The method in Figure 4-14 is called in the code on line 34, which is used to obtain a frequent item set that satisfies the minimum confidence. The method 'getCandidateCollection (itemkFcMap)' called on line 38 is used to get the degree of association between items in a frequent item set.

Line 68 in Figure 4-13 is used to find items that meet the minimum support.

Please choice options

Type
Name
ALL

Check Result

product_name	product_type	batch_number	product_supplier	outbound_time	outbound_quantity
aa	aa	A100420199935	ss	2019-04-10	2
aa	aa	A1004201991321	ss	2019-04-10	3
aa	aa	A100420191074	ss	2019-04-10	3
aa	aaaa	A1004201921554	www	2019-04-10	3
aa	aaaa	A1004201910626	www	2019-04-10	3

Figure 4-15 product outbound check result

Figure 4-15 show the product outbound check, which is implemented in the same way as the product inbound check and inventory check.

Product list

PRODUCT_NAME	PRODUCT_PRICE	PRODUCT_SUPPLIER	PRODUCT_QUANTITY	PRODUCT_STATUS	BUY
aa	23.0	ss	92	ok	• BUY
aa	2.0	www	390	ok	• BUY
apple	1.99	supplier003	130	ok	• BUY
Biscuits	1.09	OneoOOO	57	ok	• BUY
Black pepper	3.45	supplier001	248	ok	• BUY
coca	2.99	cocacola	1195	ok	• BUY
cucumber	1.99	supplier007	0	ok	• BUY
dadaa	33.0	aaa	70	ss	• BUY
dida	3.0	asy	201	ds	• BUY
daada	33.0	test	79	ok	• BUY

Showing 1 to 10 of 27 entries

Figure 4-16 product list for commodity out

Shop Cart

PRODUCT_ID	PRODUCT_NAME	PRODUCT_PRICE	QUANTITY	Total
20	aa	23.0	2	46

Showing 1 to 1 of 1 entries

Total: 46

[Check Out](#)

Figure 4-17 shopping cart

Figure 4-18 product selling information

Figures 4-16, 4-17 and 4-18 show the relevant implementation of product sales. Figure 4-16 shows the sales list of the product, which is implemented by jQuery. The dialog box of 4-17 is implemented by bootstrap. In the dialog box, after filling in the number of purchased goods, the function of automatic calculation is performed by JS + jQuery is implemented. After the user clicks the OK button in Figure 4-17, the shopping cart of Figure 4-18 is generated. The shopping cart is implemented by JS + jQuery to dynamically update the price after multiple items are added. After clicking checkout, it will be sending the HTTP Post request by Ajax, the Servlet looks for the corresponding method to process the data according to the outgoing description. The main data transmission format is JSON.

4.2 Testing

4.2.1 Unit Testing

Test Identifier	0001
Test Objective	User Login
Expected Result(s)	The database can match the username and password filled in by the user and display the dashboard
Pass/Fail	Pass

Test Identifier	0002
Test Objective	Inbound check (by name)
Expected Result(s)	The database can match the key word filled in by the user and display the check results
Pass/Fail	Pass

Test Identifier	0003
Test Objective	Inbound check (by type)
Expected Result(s)	The database can match the key word filled in by the user and display the check results
Pass/Fail	Pass

Test Identifier	0004
Test Objective	Inbound check (get all)
Expected Result(s)	The database can return all inbound details and display the check results
Pass/Fail	Pass

Test Identifier	0005
Test Objective	Add new product
Expected Result(s)	The function can insert new data into inbound table and product table
Pass/Fail	Pass

Test Identifier	0006
Test Objective	Inventory check (by name)
Expected Result(s)	The database can match the key word filled in by the user and display the check results
Pass/Fail	Pass

Test Identifier	0007
Test Objective	Inventory check (by type)
Expected Result(s)	The database can match the key word filled in by the user and display the check results
Pass/Fail	Pass

Test Identifier	0008
Test Objective	Inventory check (get all)
Expected Result(s)	The database can return all inventory details and display the check results
Pass/Fail	Pass

Test Identifier	0009
Test Objective	Inventory details update
Expected Result(s)	The database can correct update the product details base on product id
Pass/Fail	Pass

Test Identifier	0010
Test Objective	Apriori algorithm
Expected Result(s)	The Apriori algorithm can correctly get product name and association degree base on the minimum support and minimum confidence.
Pass/Fail	Pass

Test Identifier	0011
Test Objective	Sales rank
Expected Result(s)	The database can correctly obtain the quantity of products sold and sort
Pass/Fail	Pass

Test Identifier	0012
Test Objective	Commodity out
Expected Result(s)	The database can correctly update the number of products based on the number of users purchased
Pass/Fail	Pass

Test Identifier	0013
Test Objective	Outbound check (by name)
Expected Result(s)	The database can match the key word filled in by the user and display the check results
Pass/Fail	Pass

Test Identifier	0014
Test Objective	Outbound check (by type)
Expected Result(s)	The database can match the key word filled in by the user and display the check results
Pass/Fail	Pass

Test Identifier	0015
Test Objective	Outbound check (get all)
Expected Result(s)	The database can return all outbound details and display the check results
Pass/Fail	Pass

Chapter V

Conclusion

5.1 Reflection

The ultimate goal of the system is to achieve reasonable management of retail products and reasonable shelf arrangements, thereby improving product management efficiency and increasing product sales. Before studying this topic, my focus was on how to effectively improve the management efficiency of the warehouse. Therefore, I read a lot of design requirements and design requirements for inventory management software. However, I have found that such systems are usually very large and complex. So small business inventory management software has become my first direction. Then, the form of inventory management of small enterprises becomes the next problem. However, I found that there are many retail stores around, and there are shop visits every day, but they still use the original accounting method, I will turn my attention from small businesses to small retailers, ready to study how to increase the efficiency of inventory management. At that time, I was very interested in the topic of future supermarket. Amazon's self-service supermarket has also been launched, which can be said to be the prototype of the future supermarket. Therefore, for self-service supermarkets, how to manage inventory every day is a problem that needs to be solved. At this point, small retailers are the same. Therefore, for each item displayed on the shelf, we can regard it as a state in the library, and then according to the degree of correlation between the items purchased by the customer, and then put the related items together to improve the efficiency of purchase of products and inventory management. According to the analysis of shopping behavior mentioned in the second part, it can be well proved that human consumption behavior can be guided, lay a foundation for the feasibility of the system, and prove the design of the system is of great significance.

For this system, the choice of algorithm is an important problem, and it is urgent to solve it. Through the understanding of various clustering algorithms, the classic strong correlation algorithm Apriori algorithm is selected for project development.

5.2 Recommendation and future work

Due to the tight development cycle of the project, the overall project is not fully completed, and some of the functions included in the existing system can only be queried according to the given fixed value. Therefore, there are still many modifications to be implemented in this system in the future.

The first function to be added is to load the user configuration dynamically according to the logged-in user. For example, for the product inventory reminder, the system now reminds that the inventory quantity is less than 100, and the modified user can set the limit value yourself. Similarly, it also includes the minimum confidence and minimum support that the user can automatically set the algorithm.

The second point is to increase the forecast for product sales. I intend to use the time series decomposition method to achieve the whole concept and completion. However, the time is limited and only the core a priori algorithm part can be completed. A timeline decomposition method is used to predict the sales volume of the product and present it to the user in the form of a chart.

The third is the optimization of the a priori algorithm. Because the Apriori algorithm scans the database multiple times, it will generate a large number of candidate sets. To improve the applicability of the system, we must solve this problem. I found this problem. An evolutionary version of a priori algorithm called the TP-tree algorithm can be referenced.

Reference:

- [1] D. J. Hand, Heikki Mannila, and Padhraic Smyth, Principles of data mining, MIT Press, 2001
- [2] J. W. Seifert, Data Mining: an overview, CRS report for congress, 2004
- [3] Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., ... & Zhou, Z. H. (2008). Top 10 algorithms in data mining. Knowledge and information systems, 14(1), 1-37.
- [4] Zhu, H., Chen, L., Li, J., Zhou, A., Wang, P., & Wang, W. (2018, July). A General Depth-First-Search based Algorithm for Frequent Episode Discovery. In 2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD) (pp. 890-899). IEEE.
- [5] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein], Introduction to Algorithms, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7. Section 22.2: Breadth-first search, pp. 531–539.
- [6] Amine, A., Ait Mohamed, O., & Bellatreche, L. (2015). Modeling approaches and algorithms for advanced computer applications. Cham: Springer.
- [7] HU X, FENG H M, LI M W, et al. Analysis of An Enhanced Apriori Algorithms in Data Mining[J]. Netinfo Security, 2015, (11):77-83.
- [8] Richards, G. (2018). *Warehouse management*. London: Kogan Page Limited.
- [9] Botta, A., De Donato, W., Persico, V., & Pescapé, A. (2016). Integration of cloud computing and internet of things: a survey. *Future generation computer systems*, 56, 684-700.
- [10] S. Subashini, V. Kavitha A survey on security issues in service delivery models of Cloud computing J. Netw. Comput. Appl., 34 (1) (2011), pp. 1-11
- [11] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, et al. A view of Cloud computing Commun. ACM, 53 (4) (2010), pp. 50-58
- [12] Leslie, J. (2016). *Behavior Analysis*. Taylor and Francis.
- [13] shopping behavior analysis. (2019).
- [14] Carey Wodehouse. What is CSS? Cascading Style Sheets - Hiring Headquarters. Upwork Global Inc.(2018, April)
- [15] HTML5 specification finalized, squabbling over specs continues. Ars Technica. [2014-10-29].
- [16] Usage Statistics and Market Share of JavaScript Libraries for Websites, April 2019. (2019).
- [17] Flanagan, D. (2011). *JavaScript*. Beijing: O'Reilly.

[18]WebCite query result. (2019).

[19]JavaScript: The World's Most Misunderstood Programming Language. [2016-08-17].

[20] twbs/bootstrap. (2019). Retrieved from <https://github.com/twbs/bootstrap>

[21] Java Software | Oracle. (2019). Retrieved from <https://www.oracle.com/java/>

[22]Porter, B., Zyl, J., & Lamy, O. (2019). Maven – Welcome to Apache Maven. Retrieved from <https://maven.apache.org/>